



Article

# Genetic Variant Analyses Identify Novel Candidate Autism Risk Genes from a Highly Consanguineous Cohort of 104 Families from Oman

Vijay Gupta <sup>1,†</sup> , Afif Ben-Mahmoud <sup>1,†</sup>, Ahmed B. Idris <sup>2</sup> , Jouke-Jan Hottenga <sup>1,3</sup> , Wesal Habbab <sup>1</sup> ,  
Abeer Alsayegh <sup>4</sup>, Hyung-Goo Kim <sup>1,5</sup>, Watfa AL-Mamari <sup>2,\*</sup> and Lawrence W. Stanton <sup>1,6,\*</sup>

<sup>1</sup> Neurological Disorder Research Center, Qatar Biomedical Research Institute (QBRI), Hamad Bin Khalifa University (HBKU), Qatar Foundation, Doha P.O. Box 5825, Qatar; vgupta@hbku.edu.qa (V.G.); abenmahmoud@hbku.edu.qa (A.B.-M.); jjhottenga@gmail.com (J.-J.H.); whabbab@hbku.edu.qa (W.H.); hk1047@rwjms.rutgers.edu (H.-G.K.)

<sup>2</sup> Developmental Paediatric Unit, Sultan Qaboos University Hospital, Sultan Qaboos University, Muscat 123, Oman; ahmed30411@gmail.com

<sup>3</sup> Department of Biological Psychology, Vrije Universiteit Amsterdam, 1081 BT Amsterdam, The Netherlands

<sup>4</sup> Genomics Department, Sultan Qaboos Comprehensive Cancer Care and Research Center, University Medical City, Muscat 123, Oman; abeer.alsayegh@gmail.com

<sup>5</sup> Department of Neurosurgery, Robert Wood Johnson Medical School, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA

<sup>6</sup> College of Health & Life Sciences, Hamad Bin Khalifa University (HBKU), Qatar Foundation, Doha P.O. Box 5825, Qatar

\* Correspondence: watfa.almamari@gmail.com (W.A.-M.); lstanton@hbku.edu.qa (L.W.S.); Tel.: +974-445-42444 (L.W.S.)

† These authors contributed equally to this work.



**Citation:** Gupta, V.; Ben-Mahmoud, A.; Idris, A.B.; Hottenga, J.-J.; Habbab, W.; Alsayegh, A.; Kim, H.-G.; AL-Mamari, W.; Stanton, L.W. Genetic Variant Analyses Identify Novel Candidate Autism Risk Genes from a Highly Consanguineous Cohort of 104 Families from Oman. *Int. J. Mol. Sci.* **2024**, *25*, 13700. <https://doi.org/10.3390/ijms252413700>

Academic Editor: Daniele Bottai

Received: 28 November 2024

Revised: 14 December 2024

Accepted: 17 December 2024

Published: 21 December 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Deficits in social communication, restricted interests, and repetitive behaviours are hallmarks of autism spectrum disorder (ASD). Despite high genetic heritability, the majority of clinically diagnosed ASD cases have unknown genetic origins. We performed genome sequencing on mothers, fathers, and affected individuals from 104 families with ASD in Oman, a Middle Eastern country underrepresented in international genetic studies. This approach identified 48 novel candidate genes significantly associated with ASD in Oman. In particular, 35 of these genes have been previously implicated in neurodevelopmental disorders (NDDs) in other populations, underscoring the conserved genetic basis of ASD across ethnicities. Genetic variants within these candidate genes that would impact the encoded protein included 1 insertion, 4 frameshift, 6 splicing, 12 nonsense, and 67 missense changes. Notably, 61% of the SNVs were homozygous, suggesting a prominent recessive genetic architecture for ASD in this unique population. The scarcity of genetic studies on ASD in the Arabian Peninsula has impeded the understanding of the unique genetic landscape of ASD in this region. These findings help bridge this knowledge gap and provide valuable insights into the complex genetic basis of ASD in Oman.

**Keywords:** autism spectrum disorder (ASD); genome sequencing (GS); variant of unknown significance (VUS); candidate ASD risk genes

## 1. Introduction

ASD is a neurodevelopmental condition characterized by difficulties in social communication and interaction, often accompanied by restricted and repetitive behaviours, interests, or activities [1,2]. Based on the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), ASD symptoms vary in severity and include challenges in social interactions, such as difficulty with eye contact, understanding social cues, or engaging in reciprocal conversations [3–5]. Globally, ASD affects approximately 1% of children, with a pronounced sex disparity, as males are diagnosed about four times more

frequently than females [3,5]. Individuals with ASD frequently display co-occurring mental, neurological, or physical comorbidities, including intellectual disability, seizures, sleep disorders, craniofacial anomalies, and gastrointestinal problems, suggesting complex genetic etiologies [6–9].

Populations in the Middle East and North Africa (MENA) region exhibit high genetic diversity and are underrepresented in genetic studies, especially in autism research [10]. This diversity stems from factors like historical migration patterns, tribal ancestry, and high consanguinity rates, ranging from 20% to 50% across the Middle East, particularly in the Arab Gulf countries [10–14]. Such high consanguinity rates are associated with a higher prevalence of genetic disorders and birth defects [11–17]. Distinctive genomic patterns in MENA populations, particularly in Saudi Arabia, the UAE, Oman, and Qatar, provide unique insights into genetic factors underlying both common and rare diseases. These patterns may also reveal specific ASD risk genes that have remained undetected in other populations [15,18–22].

ASD research in the Middle East has grown significantly in recent years. Notably, a study in Qatar reported an ASD prevalence rate of 1 in 87 children aged 6 and 11, aligning with global rates and challenging previous assumptions of lower prevalence in this region [23]. This has raised awareness and driven further ASD research and support in the Middle East [18,24–27]. In 2019, a pioneering study found ASD prevalence in Omani children to be 15 times higher than previous estimates, highlighting the importance of proper diagnostic methods and enhanced ASD awareness [28].

ASD is associated with highly heterogeneous genetic mutations affecting various biological processes and pathways [29–31], including synaptic plasticity, chromatin remodeling, gene transcription, and protein degradation [32–36]. These mutations include single nucleotide variants (SNVs), rare copy number variations (CNVs), and chromosomal structural changes [33,35,37,38]. Traditional positional cloning and linkage analysis have identified only a limited number of ASD genes, hampered by the scarcity of families with multiple affected members and patients with chromosomal anomalies. However, the advent of next-generation sequencing (NGS) technologies such as exome sequencing (ES) and genome sequencing (GS) have revolutionized gene discovery by generating massive genomic datasets that greatly expand the scope of genetic research [19,30,31,39–52]. Simultaneously, large-cohort analyses have become increasingly prominent, revealing critical associations between genetic variants and human diseases across diverse populations [29–31,39,42]. Together, these innovations are advancing precision medicine, enabling researchers to more accurately map genetic risk factors and develop tailored interventions at the individual level.

Extensive studies highlight *de novo* variants (DNVs) as major contributors to ASD genetics, while recessive inherited variants appear less common, with estimates of 1.1% in Autism Speaks MSSNG database and 0.3% in the Simons Simplex Collection [53]. However, recessive variants are increasingly recognized as contributors to ASD, particularly in consanguineous populations, where it may account for ~5% of cases [54,55]. Intriguingly, studies of populations with high consanguinity report up to 39% of ASD cases associated with recessive inheritance, underscoring the importance of this genetic pattern in these populations [56]. Examining populations with high consanguinity is essential to comprehensively understand ASD's genetic architecture, including both *de novo* and recessive inheritance patterns [57,58].

Herein, we enrolled a cohort of 104 family trios from Oman, each consisting of an ASD-affected individual and both unaffected parents, and we utilized GS to identify *de novo* and inherited genetic variants that might contribute to disease risk. Using rigorous selection criteria, we identified 83 ASD candidate genes, including 35 genes previously associated with NDDs in global populations, as well as 48 genes not associated with ASD previously. This study provides a comprehensive genetic investigation of ASD in the Omani population, laying a critical foundation for understanding the genetic complexity of ASD in a region with high consanguinity.

## 2. Results

### 2.1. Cohort Description

The current study was approved by the Medical Research Ethics Committee of Sultan Qaboos University in Muscat and the Institutional Review Board of Qatar Biomedical Research Institute. Written informed consents were obtained from the parents, following strict adherence to the ethical tenets of the Declaration of Helsinki. A total of 312 participants, comprised of 104 probands and both parents, were recruited primarily from the clinics in the Genetics and Developmental Medicine departments at Sultan Qaboos University hospitals. Recruitment of the cohort and the sample collection took place between the years of 2015 and 2022 in the Muscat and Batinah governorates, consisting of densely populated Omani populations. All participants were Omani nationals, and all the parents were negative for autistic traits. The cohort consists of 69 males and 35 females. A total of 59 (57%) families self-reported consanguinity in their marital history, with 61% male and 39% female probands among consanguineous families. Each subject received a clinical diagnosis of ASD based on the criteria outlined in the DSM-5 by the American Psychiatric Association (2013) [4]. Potential confounding variables, including environmental exposures, birth complications, and other pregnancy-related conditions, were evaluated for possible exclusion from the study throughout the recruitment phase. Within this cohort, the most frequently observed comorbidities were language and speech delays (78%), behavioural problems (hyperactive, aggressive, and self-injurious behaviours) (60%), developmental delays (42%), intellectual disability (13%), and epilepsy (12%) (Table 1).

**Table 1.** Summary of clinical features of Omani cohort with associated comorbidities.

Sex	Number of Individuals (%)
	Male—66%
	Female—34%
Consanguinity (parents)	56%
Ethnicity	100%
Autism	100%
Language/speech delay	78%
Behavioural problems	60%
Developmental Delays	42%
Intellectual disability	13%
Epilepsy/seizures/spasms	12%

All the parents were negative for autistic traits.

### 2.2. Identification of Candidate ASD Risk Genes in the Omani ASD Cohort

We performed genome sequencing (GS) on both parents and their ASD-affected children (“trios”) from each of the recruited 104 families. The GS data were aligned to the human genome reference sequence (hg38 version), and raw Fastq files were converted to variant call format (VCF) files using the DRAGEN genomics pipeline v.4.2.4 [59]. Two types of VCF files were obtained, one for single nucleotide variants (SNVs) and another for structural and copy number variants (CNVs). For SNVs, VCF files were analysed using QCI Interpret by Qiagen (<https://apps.ingenuity.com/qci>) (accessed on 10 March 2024) to identify and prioritize genomic variants associated with ASD. The final GS dataset analysis workflow began with VCF, followed by filtering variants based on multiple criteria, including sequence quality, predicted deleteriousness of the encoded protein, and presence of the variant in common databases such as gnomAD, ExAC (>0.1%). Key filtering criteria to prioritize variants with a predicted deleterious effect on protein function included a high (>20) Combined Annotation Dependent Deletion (CADD) score. The pathogenicity of the identified variants was assessed using guidelines from the American College of

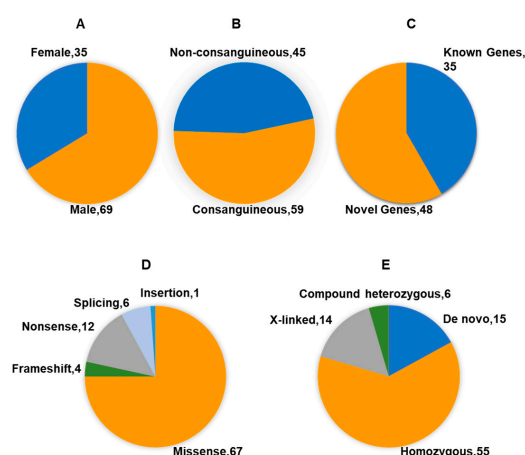
Medical Genetics and Genomics (ACMG), along with gnomAD frequency, CADD scores, and various other bioinformatic predictions. We focused on variants classified by ACMG as “variants of unknown significance” (VUS), “likely pathogenic”, or “pathogenic”. Variants classified as “benign” or “likely benign” were excluded from further analysis.

We hypothesized that the probands would carry either known or novel risk gene variants, arising de novo or inherited in autosomal dominant heterozygous, autosomal recessive homozygous, or X-linked recessive hemizygous patterns. Several bioinformatics software packages were used to assess the conservation of these variants and estimate their potential negative impact on gene function. Variants with a minor allele frequency of less than 0.1% that resulted in distinct protein-altering changes (nonsense, insertions or deletions, missense, and splice sites) were prioritized using various protein prediction software (SIFT, PolyPhen-2, CADD, and M-CAP).

### 2.3. Refinement and Prioritization of ASD Candidate Genes

Initially, we obtained 126 variants across 116 different genes from the analysis of 104 trio datasets (Table 2). To further refine these findings, we incorporated gene loss tolerance metrics using calculated Z-scores and pLI scores (indicating the likelihood of being intolerant to loss-of-function). For de novo variants, particularly nonsense/frameshift variants, we applied a strict pLI score threshold greater than 0.9; missense variants were prioritized using a dual criterion: a Z-score above 3.0 combined with a pLI score greater than 0.9. Variants in ASD candidate genes that did not meet these criteria were excluded from further consideration. pLI and Z-scores are primarily developed for evaluating heterozygous variants and the impact of haploinsufficiency. Genes associated with dominant disorders tend to have higher pLI scores because even one functional copy of the gene being disrupted can cause a disorder. In contrast, recessive genes can tolerate LoF mutations in one allele because the other functional copy compensates, leading to lower pLI scores. Homozygous variants in genes with low pLI scores are often mistakenly considered benign, yet they can cause severe disease in recessive disorders. Our final analysis revealed 90 SNVs from 83 unique candidate genes (Table 2).

The biological and clinical relevance of these genes and their pathways was evaluated through a literature review and by considering the participants’ clinical features. The identified variants included 1 insertion, 4 frameshift, 6 splicing, 12 nonsense, and 67 missense changes. Based on the inheritance pattern, we identified 15 de novo, 55 homozygous, 6 compound heterozygous, and 14 X-linked variants distributed across 35 female and 69 male probands (Table 2 and Figure 1). Detailed criteria and supporting evidence for each variant are provided in Table 2.



**Figure 1.** Summary of 104 ASD individuals and their GS results. (A) Sex distribution of participants; (B) number of consanguineous families; (C) counts of known and novel genes identified; (D) types of genetic variants identified; (E) inheritance patterns of the variants.

**Table 2.** Overview of single nucleotide variants identified from genome sequencing (GS) analysis of 104 Omani autism individuals.

Subject	Gender	Consanguinity	Genes	Accession Number	Chr.	Nucleotide Change	Protein Change	Type of Variant	Inheritance (GENOTYPE)	gnomAD (AF)	QGP (AF)	CADD	ACMG Interpretation	Z Score	pLI Score	Previously Linked with ASD/NDDs
1	M	+	<i>HNRNPCL4</i>	NM_001302551.2	1	c.830C>T	p.A277V	missense	homoz.	0	0	21.2	LP	NA	NA	Novel
2	M	-	<i>PNCK</i>	NM_001366977.1	X	c.932G>A	p.R311Q	missense	x-linked	0	0.000238598	21.1	LP	NA	NA	Novel
3	M	-	<i>COP1</i>	NM_022457.7	1	c.1291C>T	p.R431W	missense	de novo	0	0	25.8	P	-0.36	0.92	Novel
			<i>TRIM73</i>	NM_198924.4	7	c.487C>T	p.R163*	nonsense	homoz.	0	0	35	P	3.23	0.93	Novel
4	M	+	<i>TSKU</i>	NM_015516.4	11	c.979C>T	p.R327W	missense	homoz.	0	0	25	LP	1.22	0.17	Novel
5	F	-	<i>SLC6A1</i>	NM_003042.4	3	c.1328G>T	p.G443V	missense	de novo	0	0	27.9	P	4.92	1	Known
			<i>CAPN8</i>	NM_001143962.2	1	c.730-2A>G	splicing	splicing	de novo	0	0	25.9	P	-0.19	0	Novel
			<i>MYO7A</i>	NM_000260.4	11	c.2489G>A	p.R830H	missense	homoz.	0.0114, European	0.00214739	32	VUS	0.06	0	Known
6	M	-	<i>RAB11FIP3</i>	NM_014700.4	16	c.1396-2A>G	splicing	splicing	de novo	0	0	35	P	-0.86	1	Novel
7	M	+	<i>XYLT1</i>	NM_022166.4	16	c.1129C>G	p.Q377E	missense	homoz.	0	0.0144863	22.3	LP	0.75	0.99	Known
8	F	-	<i>SYTL1</i>	NM_001193308.2	1	c.1317_1323-delAGGATCC	p.G440fs*?	frameshift	de novo	0	0		P	-0.52	0	Novel
9	M	-	<i>LAS1L</i>	NM_031206.7	X	c.1082C>G	p.P361R	missense	X-linked	0	0.00572636	22.1	LP	NA	NA	Known
10	M	-	<i>RIBC1</i>	NM_001031745.5	X	c.106C>G	p.R36G	missense	X-linked	0	0	23.9	LP	NA	NA	Novel
11	F	+	<i>KIAA1755</i>	NM_001029864.2	20	c.2052+1G>A	splicing	splicing	homoz.	0.0024, African	0.00105665	33	P	-0.24	0	Novel
12	M	-	<i>EDEM1</i>	NM_014674.3	3	c.1301G>A	p.W434*	nonsense	de novo	0	0	45	P	-1.39	0	Novel
			<i>SLC12A9</i>	NM_020246.4	7	c.764A>C	p.Y255S	missense	de novo	0	0	28.5	LP	1.96	0	Novel
13	M	-	<i>AMOT</i>	NM_001113490.2	X	c.514C>T	p.R172C	missense	X-linked	0	0	29	LP	NA	NA	Novel
14	F	+	<i>VWA8</i>	NM_015058.2	13	c.5093G>A	p.R1698Q	missense	homoz.	0.0038, European	$3.40855 \times 10^{-5}$	29.9	VUS	0.51	0	Novel
			<i>H6PD</i>	NM_004285.4	1	c.127T>C	p.W43R	missense	homoz.	0.0013, European	$3.40855 \times 10^{-5}$	29.6	VUS	0.11	0	Novel
			<i>KIAA1755</i>	NM_001029864.2	20	c.19G>A	p.D7N	missense	homoz.	0	0.000340855	25.3	LP	-0.24	0	Novel
15	M	+	<i>TTC36</i>	NM_001080441.4	11	c.284G>C	p.R95P	missense	homoz.	0	0	23.6	LP	-1.92	0	Novel
16	M	-	<i>LILRB3</i>	NM_006864.4	19	c.1050T>G	p.Y350*	nonsense	homoz.	0	0	23.2	P	1.45	0	Novel
			<i>TMEM31</i>	NM_182541.2	X	c.142C>A	p.Q48K	missense	X-linked	0.0094, Latino	0.000443111	21.7	VUS	NA	NA	Novel
17	M	+	<i>PRRT2</i>	NM_145239.3	16	c.649dupC	p.R217fs*8	frameshift	de novo	0	0	31	P	-0.26	0.65	Known
18	F	+	<i>TRIM15</i>	NM_033229.3	6	c.608A>G	p.D203G	missense	de novo	0	0	23.5	LP	-0.18	0	Novel
			<i>LTBP1</i>	NM_206943.4	2	c.3616T>C	p.C1206R	missense	homoz.	0	0	29.8	LP	-0.7	0	Novel

Table 2. Cont.

Subject	Gender	Consanguinity	Genes	Accession Number	Chr.	Nucleotide Change	Protein Change	Type of Variant	Inheritance (GENOTYPE)	gnomAD (AF)	QGP (AF)	CADD	ACMG Interpretation	Z Score	pLI Score	Previously Linked with ASD/NDDs	
19	M	-	CNV identified														
20	M	+	<i>ATP13A5</i>	NM_198505.4	3	c.632A>G	p.Q211R	missense	de novo	0	0	26.1	LP	1.24	0	Novel	
			<i>COL6A6</i>	NM_001102608.3	3	c.5002G>T	p.G1668*	nonsense	homoz.	0	0.000340855	47	P	0.53	0	Novel	
21	M	-	<i>RORC</i>	NM_005060.4	1	c.253C>T	p.H85Y	missense	homoz.	0.0654	0.0116572	26.9	VUS	2.89	1	Novel	
22	M	+	<i>ERBB4</i>	NM_005235.3	2	c.2900T>G	p.F967C	missense	homoz.	0	0.000204513	31	LP	2.82	1	Known	
			<i>FMR1</i>	NM_002024.6	X	c.716C>T	p.A239V	missense	X-linked	0	0	25.3	LP	NA	NA	Known	
23	M	+	<i>ECH1</i>	NM_001398.3	19	c.284A>G	p.K95R	missense	homoz.	0	0	26	LP	-0.17	0	Novel	
			<i>ZNF250</i>	NM_001109689.4	8	c.593A>T	p.E198V	missense	de novo	0	0	23.4	LP	3.41	0.93	Novel	
24	M	+	<i>AGTR2</i>	NM_000686.5	X	c.411C>A	p.C137*	nonsense	X-linked	0	$3.40855 \times 10^{-5}$	31	P	NA	NA	Known	
25	F	-	CNV identified														
26	F	-	<i>DENND2A</i>	NM_015689.5	7	c.2971G>A	p.G991S	missense	homoz.	0.0024, African	0.000204513	25.8	LP	1.84	0	Novel	
27	M	-	<i>HTT</i>	NM_001388492.1	4	c.102_110dup-GCAGCAGCA	p.Q36_Q38dup	in-frame Insertion	de novo	0	0	NA	LP	3.7	1	Known	
28	F	-	<i>HIVEP3</i>	NM_024503.5	1	c.3316C>T	p.Q1106*	nonsense	de novo	0	0	35	P	NA	NA	Novel	
			<i>PLIN4</i>	NM_001367868.2	19	c.2246_2247insG	p.G750fs*14	frameshift	de novo	0.0151, Latino	0	23.9	P	-1.44	0	Novel	
29	M		<i>RAB3B</i>	NM_002867.4	1	c.8C>T	p.S3L	missense	de novo	0	0	23.7	LP	NA	NA	Novel	
			<i>PGAP3</i>	NM_033419.5	17	c.717G>C	p.W239C	missense	homoz.	0.0048, African	0	31	VUS	-0.47	0	Novel	
30	M	+	<i>LMF1</i>	NM_022773.4	16	c.529C>T	p.L177F	missense	homoz.	0	0.00146568	28.2	LP	-2.56	0	Known	
					17	c.12650T>C	p.I4217T	missense		0	0	26.9	LP	-2.01	0		
31	M	-	<i>DNAH17</i>	NM_173628.4	17	c.2121C>A	p.N707K	missense	comp. het.	0	$3.40855 \times 10^{-5}$	10	LP			Known	
32	M	-	<i>ACAP9</i>	NM_001190810.1	10	c.1930_1931ins-TCTGGTA	p.T644fs*?	frameshift	homoz.	0	0	NA	P	4.15	0	Novel	
			<i>OBSCN</i>	NM_001386125.1	1	c.2186A>C	p.E729A	missense	homoz.	0	0.00654441	25	LP	-0.82	0	Known	
33	M	+	<i>HAS2</i>	NM_005328.3	8	c.221C>G	p.S74C	missense	homoz.	0	0	23.4	LP	4.42	1	Novel	
			<i>ZNF81</i>	NM_007137.5	X	c.980A>G	p.E327G	missense	X-linked	0.0188, Latino	0.000579493	24.2	VUS	NA	NA	Known	

Table 2. Cont.

Subject	Gender	Consanguinity	Genes	Accession Number	Chr.	Nucleotide Change	Protein Change	Type of Variant	Inheritance (GENOTYPE)	gnomAD (AF)	QGP (AF)	CADD	ACMG Interpretation	Z Score	pLI Score	Previously Linked with ASD/NDDs
34	M	-	MAP3K9	NM_001284230.2	14	c.1790G>A	p.R597Q	missense	de novo	0	$3.40855 \times 10^{-5}$	23.8	LP	1.6	0	Novel
			MAGEC3	NM_138702.1	X	c.212C>A	p.S71*	nonsense	X-linked	0.0017 European	0	26.3	P	NA	NA	Novel
35	F	-	ABCA13	NM_152701.5	7	c.633-2A>G	splicing	splicing	de novo	0	0	33	P	-0.12	0	Known
			CLEC18C	NM_173619.4	16	c.994G>A	p.G332R	missense	homoz.	0	0	29	LP	3.26	0.94	Novel
36	M	+	ZNF483	NM_133464.5	9	c.2047C>T	p.R683*	nonsense	homoz.	0	0	35	P	3	0.97	Novel
			PREP	NM_002726.5	6	c.1213+1G>A	splicing	splicing	homoz.	0.0013 European	$3.40855 \times 10^{-5}$	34	P	2.65	1	Novel
37	M	+	ZNF207	NM_001098507.2	17	c.637G>C	p.G213R	missense	homoz.	0	0	25	LP	2.69	0.12	Novel
			CARMIL1	NM_017640.6	6	c.3149A>T	p.H1050L	missense	homoz.	0	0.00132933	20.7	LP	2.03	0	Novel
38	F	NA	KCNMA1	NM_001161352.2	10	c.2725C>T	p.R909W	missense	de novo	0	0	32	LP	6.52	1	Known
			TMEM184A	NM_001097620.2	7	c.365C>T	p.S122F	missense	de novo	0	0	22.8	LP	-0.75	0	Novel
39	M	-	CRYBG1	NM_001371242.2	6	c.4916A>T	p.K1639I	missense	homoz.	0	0	22.5	LP	0.93	0	Novel
			NUS1	NM_138459.5	6	c.684T>A	p.S228R	missense	homoz.	0	0	22	LP	1.03	1	Known
40	M	+	KCTD6	NM_001128214.2	3	c.538G>A	p.G180R	missense	homoz.	0	0	23.1	LP	2.86	0.38	Novel
41	F	+	PDE2A	NM_002599.5	11	c.229C>T	p.R77*	nonsense	homoz.	0	0	34	P	3.86	0.02	Known
42	M	-	ARHGAP33	NM_001366178.1	19	c.3691C>T	p.P1231S	missense	homoz.	0	0.000272684	21.4	LP	1.78	0	Novel
43	F	+	UBE2Q2	NM_173469.4	15	c.1031A>T	p.N344I	missense	homoz.	0.0065	0.000715795	25	VUS	0.38	0	Novel
44	M	+	SPATA31A3	NM_001083124.1	9	c.1879C>T	p.Q627*	nonsense	homoz.	0	0	25.9	P	5.51	0.65	Novel
45	M	+	SORCS2	NM_020777.3	4	c.1662G>A	p.W554*	nonsense	homoz.	0	0	47	P	-1.75	0	Novel
46	M	+	CWF19L2	NM_152434.3	11	c.270G>T	p.K90N	missense	de novo	0	$3.40994 \times 10^{-5}$	23	LP	-0.65	0	Novel
			CYB5B	NM_030579.3	16	c.212G>A	p.G71D	missense	homoz.	0	0	28.5	LP	0.81	0.01	Novel
47	F	+	HSPG2	NM_005529.7	1	c.1369A>G	p.S457G	missense	homoz.	0	0	27.9	LP	3.4	0	Known
			SPTA1	NM_003126.4	1	c.2195T>A	p.L732Q	missense	de novo	0	0	25.4	LP	-0.8	0	Known
48	M	+	SCN2A	NM_001040142.2	2	c.3932T>G	p.L1311R	missense	de novo	0	0	29.3	LP	8.71	1	Known
49	F	-	NOTCH4	NM_004557.4	6	c.1522C>T	p.Q508*	nonsense	de novo	0	0	38	P	2.17	0	Novel
			MTUS1	NM_001363059.2	8	c.1165C>T	p.Q389*	nonsense	de novo	0	0.00149976	36	P	-5.75	0	Novel
50	F	+	NHLRC1	NM_198586.3	6	c.946T>A	p.F316I	missense	de novo	0	0	22.4	LP	-0.09	0	Known
51	M	+														
52	F	-	CNV identified													

Table 2. Cont.

Subject	Gender	Consanguinity	Genes	Accession Number	Chr.	Nucleotide Change	Protein Change	Type of Variant	Inheritance (GENOTYPE)	gnomAD (AF)	QGP (AF)	CADD	ACMG Interpretation	Z Score	pLI Score	Previously Linked with ASD/NDDs	
53	M	+	<i>SEC61A1</i>	NM_013336.4	3	c.331C>G	p.L111V	missense	homoz.	0	0	23.4	LP	4.94	1	Novel	
54	F	+	CNV identified														
55	M	+	<i>RNF113A</i>	NM_006978.3	X	c.76G>A	p.G26R	missense	X-linked	0	0.00037494	21.7	LP	NA	NA	Novel	
56	M	-	<i>TNRC18</i>	NM_001080495.3	7	c.487+126delT	Intronic	intronic	de novo	0	0	NA	LP	-7.54	0	Novel	
57	M	-	<i>PHF14</i>	NM_001007157.2	7	c.73A>C	p.S25R	missense	de novo	0	0	29.9	LP	0.87	0.03	Known	
58	M	-	<i>SPATA31A3</i>	NM_001083124.1	9	c.1879C>T	p.Q627*	nonsense	homoz.	0	0	25.9	LP	5.51	0.65	Novel	
59	M	+	<i>CCT6A</i>	NM_001762.4	7	c.336+1G>A	Splicing	splicing	de novo	0	$3.40855 \times 10^{-5}$	34	P	1.42	1	Novel	
60	F	-	<i>ESYT3</i>	NM_031913.5	3	c.2468+1G>T	Splicing	splicing	homoz.	0	0.00528325	34	P	0.17	0	Novel	
61	M	+	<i>UBE3C</i>	NM_014671.3	7	c.916_917delAG	p.S306*	frameshift	homoz.	0	0	NA	P	3.04	1	Novel	
62	F	+	<i>MANSC1</i>	NM_018050.4	12	c.250T>G	p.F84V	missense	comp. het.	0	0	27.3	LP	0.82	0	Novel	
					12	c.38T>A	p.L13*	nonsense				0.00037494	34	P	0.82		0
63	F	+	<i>SLC19A1</i>	NM_194255.4	21	c.1514A>G	p.E505G	missense	homoz.	0	0.00109074	22.9	LP	-0.71	0	Novel	
64	M	-	<i>LRP1B</i>	NM_018557.3	2	c.1560G>T	p.K520N	missense	homoz.	0	0	26.3	LP	3.33	1	Known	
65	M	-	<i>VIRMA</i>	NM_015496.5	8	c.1069A>T	p.T357S	missense	de novo	0	$3.40855 \times 10^{-5}$	25.2	LP	3.68	1	Novel	
					11	c.973T>C	p.F325L	missense	de novo			0	0	24.2	LP		1.13
66	M	+	<i>DNAH17</i>	NM_173628.4	17	c.12389C>T	p.P4130L	missense	comp. het.	0	0.000102256	27.5	LP	-2.01	0	Known	
					17	c.11825G>A	p.R3942Q	missense				1.4	0.00453337	23.5	LP		-2.01
67	M	+	<i>RNF175</i>	NM_173662.4	4	c.247-1G>A	splicing	splicing	homoz.	0	$3.40855 \times 10^{-5}$	33	P	0.45	0	Novel	
68	M	-	<i>PLBD2</i>	NM_173542.4	12	c.1576C>T	p.R526C	missense	comp. het.	0.05	0.000920308	25.8	VUS	-0.17	0	Novel	
					12	c.1012C>T	p.R338W	missense				0.07	0.000272684	24.6	VUS		-0.17
69	F	+	<i>DNAH3</i>	NM_001347886.2	16	c.2023G>T	p.V675F	missense	homoz.	0	0.0058627	23	LP	2.47	0	Known	
70	M	-	<i>RECQL4</i>	NM_004260.4	8	c.2386G>A	p.E796K	missense	comp. het.	0.01	$3.40855 \times 10^{-5}$	25.6	VUS	-6.29	0	Known	
					8	c.3501C>T	p.I1167I	missense				0.01	$6.8171 \times 10^{-5}$	NA	VUS		-6.29
71	M	-	<i>UBR4</i>	NM_020765.3	1	c.2551G>A	p.V851M	missense	comp. het.	0.02	0.000852137	23	VUS	8.42	1	Known	
					1	c.3137G>A	p.R1046Q	missense				0.0024	0.000136342	23.5	VUS		8.42
72	F	+	<i>SNX21</i>	NM_033421.4	20	c.287C>T	p.A96V	missense	de novo	0	0	26	LP	-0.17	0	Novel	





Table 2. Cont.

Subject	Gender	Consanguinity	Genes	Accession Number	Chr.	Nucleotide Change	Protein Change	Type of Variant	Inheritance (GENOTYPE)	gnomAD (AF)	QGP (AF)	CADD	ACMG Interpretation	Z Score	pLI Score	Previously Linked with ASD/NDDs
94	M	+	<i>TMEM259</i>	NM_001033026.2	19	c.380A>G	p.Q127R	missense	de novo	0	0	25	LP	−2.09	0	Novel
95	M	-	<i>INTS6L</i>	NM_001351601.3	X	c.2584G>A	p.A862T	missense	X-linked	0	$6.8171 \times 10^{-5}$	22.2	LP	NA	NA	Novel
96	F	-	<i>GIPR</i>	NM_000164.4	19	c.1152G>T	p.Q384H	missense	comp. het.	0	0.000409026	36	LP	−1.18	0	Novel
					19	c.302G>A	p.R101H	missense		1.46	0.000647624	25.6	VUS	−1.18	0	
97	F	+	<i>SLC16A5</i>	NM_004695.4	17	c.380C>G	p.T127R	missense	de novo	0	0	23.9	LP	1.33	0	Novel
98	F	+	CNV identified													
99	M	+	<i>ARHGAP4</i>	NM_001666.5	X	c.1339T>C	p.Y447H	missense	X-linked	0	$6.8171 \times 10^{-5}$	24.9	LP	NA	NA	Novel
100	M	-	<i>GIPR</i>	NM_000164.4	19	c.1264C>T	p.Q422*	nonsense	de novo	0	0.000749932	36	P	−1.18	0	Novel
			<i>MAP3K5</i>	NM_005923.4	6	c.439T>C	p.Y147H	missense	de novo	0	0	32	LP	3.26	1	Novel
			<i>CIP2A</i>	NM_020890.3	3	c.2303C>G	p.S768*	nonsense	de novo	0	0	41	P	−0.62	0	Novel
101	M	+	<i>CLCNKA</i>	NM_004070.4	1	c.781+1G>A	splicing	splicing	de novo	0	$3.40855 \times 10^{-5}$	34	P	1.05	0	Known
102	M	+	<i>SAFB2</i>	NM_014649.3	19	c.1798G>A	p.D600N	missense	de novo	0	0	28.5	LP	−0.14	0.99	Novel
103	F	-	<i>ZNF827</i>	NM_001306215.2	4	c.1892A>T	p.D631V	missense	de novo	0	0	23.1	LP	2.95	1	Novel
104	F	+	<i>DENR</i>	NM_003677.5	12	c.434A>G	p.Q145R	missense	de novo	0	0	22.7	LP	1.6	0.92	Novel

A total of 116 unique genes were identified in 95 families, while 10 CNVs were detected in eight families. In family 93, neither SNV nor CNV was detected, and in 8 families, CNVs were detected, as shown in Table 3. Variants are denoted by their respective GenBank accession numbers. The term “NA” is used to signify “not available” when information on specific variants is absent. Of the 116 unique genes, 83 (marked in red) are considered highly likely to be causative for ASD based on criteria such as low prevalence, high-quality coverage, high CADD scores, significant pLI and Z-scores, and alignment with ACMG guidelines. These 83 genes were identified in 76 families, 6 genes found in 2 families each, and 1 gene identified in 3 families. Among the 83 genes highlighted in red, 48 are newly identified (novel), while 35 are known genes already associated with ASD or other NDDs. In the 56 consanguineous families, putative candidate genes were identified in 41 families. These included 41 homozygous recessive, 6 X-linked, and 6 de novo variants, with recessive genes accounting for 84% of the total. In the remaining 48 non-consanguineous families, we found 20 homozygous recessive, 7 X-linked, 10 de novo genes, with recessive genes comprising 56% of the findings. For de novo candidate genes, prioritization used to short-list was based on pLI scores above 0.9 and Z-score above 3. Additional metrics used include AF (allele frequency), pLI score, Z-score, and CADD score. QGP is Qatar Genome Program having allele frequencies from a control dataset of Qatari population. Chr. is chromosome and homoz. is homozygous recessive variant. Variants were classified using ACMG guidelines, which categorize genetic variants into benign, likely benign, uncertain significance (VUS), likely pathogenic (LP), and pathogenic (P) to assess their clinical significance.

**Table 3.** Ten heterozygous intragenic microCNVs identified in eight Omani ASD individuals.

Subject ID	Subject 19	Subject 25	Subject 52	Subject 54	Subject 74	Subject 75	Subject 89	Subject 98		
Sex	M	F	F	F	M	M	F	F		
CNV and genomic coordinates (hg38)	Chr16:6864585-6865758	Chr3:192671334-192672351	Chr19:4559368-4562532	Chr3:133414146-133418344	Chr10:54921983-54925066	Chr9:71289426-71290609	Chr17:45440911-45508556	Chr2:1521403-1522564	Chr10:26709200-26711140	Chr12:21165423-21166430
Type of CNV	heterozygous microdeletion	heterozygous microdeletion	heterozygous microdeletion	heterozygous microdeletion	heterozygous microdeletion	heterozygous microduplication	heterozygous microdeletion	heterozygous microdeletion	heterozygous microdeletion	heterozygous microdeletion
Size of CNV (bp)	1173	1017	3164	−4198	3083	1183	67645	1161	1940	1007
Cytoband	16p13.3	3q29	19p13.3	3q22.1	10q21.1	9q21.12	17q21.31	2p25.3	10p12.1	12p12.1
Genes involved	<i>RBFOX1</i> (NM_001142333)	<i>FGF12</i> (NM_001377292)	<i>SEMA6B</i> (NM_032108)	<i>BFSP2</i> (NM_003571)	<i>PCDH15</i> (NM_001354404)	<i>TRPM3</i> (NM_001366141)	<i>PLEKHM1</i> (NM_014798)	<i>TPO</i> (NM_000547)	<i>PDSS1</i> (NM_001321978)	<i>SLCO1B1</i> (NM_006446)
Location	intron 3	intron 2	exon 1 and intron 1	intron 1	intron 3	intron 1	exon 6 (LRRC37A4P) exon 1-intron9 (PLEKHM1)	Intron 15	Intron 4, exon 5, intron 5	intron 2
Inheritance	de novo	de novo	de novo	de novo	de novo	de novo	de novo	de novo	de novo	de novo
Disease linked	autism susceptibility 1; epilepsy, HGMD-71	developmental and epileptic encephalopathy, MIM 601513, HGMD-26	epilepsy, progressive myoclonic, MIM-608873, HGMD-43	Cataract, MIM 603212, HGMD-13	Usher syndrome type 1; nonsyndromic genetic hearing loss, MIM 605514, HGMD-225	autosomal dominant non-syndromic intellectual disability, HGMD-24	osteopetrosis, MIM- 611466, PLEKHM1 HGMD-15	familial thyroid dysmorphogenesis, MIM 600044, HGMD-264	deafness–encephaloneuropathy–obesity-valvulopathy syndrome, MIM 607429, HGMD-19	rotor syndrome, MIM 604843, HGMD-51

Types and sizes of structural variants identified in Omani ASD individuals, along with their genomic coordinates (GRCh38/hg38) and associated genes, are detailed. These unbalanced genomic CNVs encompass at least one well-established gene linked to NDDs, supporting their potential pathogenicity. HGMD-71, 26, 43, 13, 225, 24, 15, 264, 19, and 51 are the number of variants reported in HGMD database on 27 November 2024 in the above-mentioned genes, respectively.

#### 2.4. Identification of Copy Number Variants (CNV) in the Omani ASD Cohort

As previously described, the original GS data in FASTAQ files were converted to VCF files using the DRAGEN genomics pipeline. CNV detection was then performed using DRAGEN CNV workflow, which employs the GATK Germline CNV caller, a read-depth and junction-based workflow for identifying CNVs > 50 bp [59,60]. We focused exclusively on de novo CNVs. We confidently identified de novo CNVs in eight probands from families 19, 25, 52, 54, 74, 75, 89, and 98, exhibiting heterozygous microdeletions or microduplications ranging from 1 kb to 68 kb (Table 3). Subject 19 displayed an 1173 bp deletion spanning intron-3 to the 5'-UTR of *RBFOX1*, a well-known ASD gene [61–64]. Subjects 25 and 52 had deletions involving *FGF12* and *SEMA6B*, two known genes associated with ASD and epilepsy [65–68]. However, these two cases did not report epilepsy. Subjects 54, 74, 89, and 98 had CNVs in other previously reported neurodevelopmental genes, including *BFSP2*, *PCDH15*, *TRPM3*, *TPO*, *PDSS1*, and *SLCO1B1* [69–74]. Collectively, ten de novo CNVs identified in eight ASD individuals are either intragenic or affect one end of a gene, involving a total of ten candidate genes. Notably, several sporadic variants in these genes have been documented in patients with ASD, intellectual disability (ID), developmental disorders, or other NDDs in the literature, further supporting their role as ASD candidate genes [69–74] (Table 3).

#### 2.5. Identification of 35 Known NDD-Associated Genes

To further validate our results, we compared our newly identified gene list of 83 genes with the established neurodevelopmental disorder (NDD)/ASD gene databases, including the Genomics England NDD/Autism panel (<https://panelapp.genomicsengland.co.uk/panels/285>) (accessed on 1 August 2024), the Simons Foundation Autism Research Initiative (SFARI) gene list (<https://gene.sfari.org/database/human-gene>) (accessed on 1 August 2024), and The Human Gene Mutation Database (HGMD) (<https://apps.ingenuity.com>) (accessed on 1 August 2024). Genes from our list that were also present in these established databases were marked as present (“Known”) or not present “Novel” in Table 2. This comparison revealed that only 3 out of the 83 genes overlapped with the SFARI genes (*LAS1L*, *PHF14*, and *SLC6A1*), and only 7 genes matched the Genomics England NDD/Autism panel genes (*HTT*, *LAS1L*, *MYO7A*, *PHF14*, *PRRT2*, *SLC6A1*, *XYLT1*). In a reverse approach, 35 of our genes, which showed mutations across multiple subjects, were found in individuals with various NDDs such as ASD, ID, macrocephaly, and other developmental disorders, as well as neurodegenerative disorders, including Huntington’s disease, Parkinson’s disease, and ALS according to the HGMD database (Table 2). This overlap corroborates the relevance of these genes in ASD, suggesting they may play a key role in the disorder since they are already known to cause various neurodevelopmental and/or neurodegenerative disorders (Table 2).

#### 2.6. Identification of 48 Novel ASD-Risk Genes

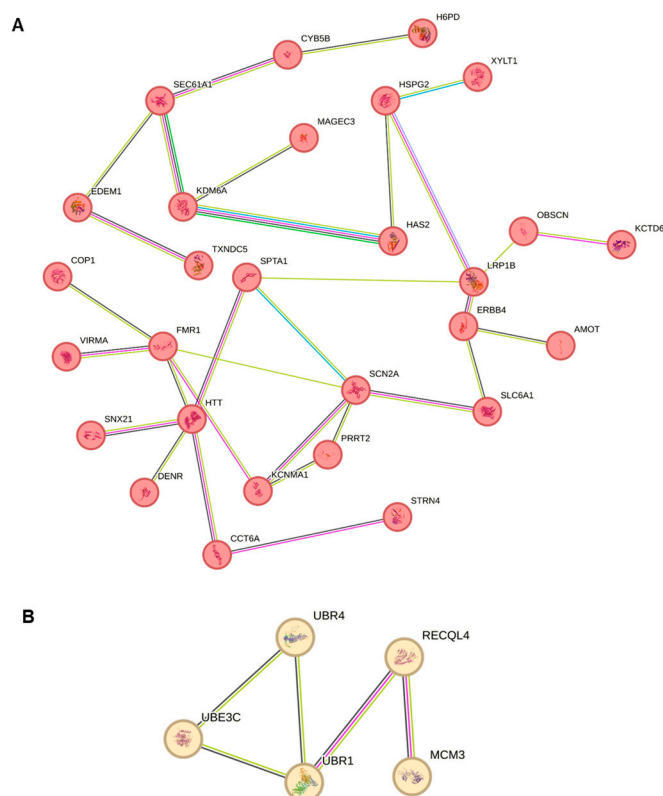
Many genes associated with NDDs may also play a role in ASD [75,76]. In this study, we employed “parent-offspring trio genome sequencing” to analyse 104 subjects with ASD, identifying 83 genes with a strong likelihood of contributing to ASD. We followed ACMG guidelines and applied rigorous criteria, including low population prevalence; high-quality sequencing coverage; elevated CADD scores, pLI, and Z-scores to prioritize and finalize the candidate gene list. As mentioned above, 35 of these genes were previously associated with NDDs or ASD, which we designate as “Known” (Table 2). In addition, we also identified 48 candidate genes not previously associated with ASD, which we refer to hereafter as “Novel” candidate genes. These novel candidate genes have limited or no information regarding their involvement in ASD, and their specific roles in ASD phenotype have not yet been described in the literature (Table 2). These novel genes contribute to a variety of biological processes, such as protein–protein and protein–DNA interactions, chromatin or DNA binding, and GTPase-activating functions, highlighting the complex genetic basis of ASD.

### 2.7. Sporadic Variants Identified in Our Novel ASD Candidate Genes

Advances in NGS have generated vast genome data from various diseases, including diabetes, cancer, cardiovascular disorders, neurodegenerative disorders, and NDDs such as ASD [32,38,45,77–88]. These genomic data provide a unique opportunity to characterize the biological functions of genes and elucidate the molecular mechanisms underlying these developmental disorders. HGMD is a comprehensive repository that catalogues genetic variants implicated in human disorders. It includes information on mutations with confirmed roles in disease causation, as well as sporadic variants found in potential disease genes reported in the literature [89]. We combed through the HGMD against our list of newly identified novel ASD candidate genes and found that most of these genes had genomic variants, in the form of SNVs or CNVs, previously linked to NDDs and neuro-degenerative disorders [29–31,38]. This raises the possibility that these genes could potentially serve as ASD candidate genes, even though they have not been formally classified as ASD-linked or ASD-susceptibility genes and are not included in SFARI or NDD/ID gene panels. Given the sporadic variants reported for these genes in HGMD, it is likely that they could be involved in disease causation.

### 2.8. Protein–Protein Enrichment Analysis of Candidate Genes

We performed protein–protein interaction enrichment analysis using STRING 3.0 (<https://string-db.org/>) (accessed on 1 September 2024) on our candidate genes (Supplementary Figure S1). Upon k-means clustering, this analysis revealed two predominant networks among the candidate genes identified (Figure 2A,B). The first network consisted of 28 nodes and 31 edges (average node degree: 2.21, average local clustering coefficient: 0.463, and PPI enrichment  $p$  value:  $5.23 \times 10^{-10}$ ) consisting of 28 proteins associated with components of the presynaptic and postsynaptic membranes. These genes include *AMOT*, *CCT6A*, *COP1*, *CYB5B*, *DENR*, *EDEM1*, *ERBB4*, *FMR1*, *H6PD*, *HAS2*, *HSPG2*, *HTT*, *KCNMA1*, *KCTD6*, *KDM6A*, *LRP1B*, *MAGEC3*, *OBSCN*, *PRRT2*, *SCN2A*, *SEC61A1*, *SLC6A1*, *SNX21*, *SPTA1*, *STRN4*, *TXNDC5*, *VIRMA*, and *XYLT1* (Figure 2A). The presynaptic membrane is essential for neurotransmitter release and synaptic transmission, processes fundamental to brain function. In ASD, abnormalities in presynaptic mechanisms have been implicated as contributors to the neurobiological features [90–92]. The second network consisted of 5 nodes and 5 edges (average node degree: 2, average local clustering coefficient: 0.667, and PPI enrichment  $p$  value:  $2.71 \times 10^{-6}$ ) and was characterized by ubiquitin-conjugating enzyme activities involving genes such as *UBR1*, *UBR4*, *UBE3C*, *RECQL4*, and *MCM3* (Figure 2B). Ubiquitin-conjugating enzymes play a vital role in the ubiquitin–proteasome system, which regulates protein degradation and turnover, thereby maintaining cellular homeostasis and various cellular functions. With respect to ASD, disruptions in ubiquitin-conjugating enzyme activity could affect protein homeostasis, synaptic function, inflammation, stress responses, as well as neuronal differentiation, migration, and synaptogenesis, which are all relevant to neurodevelopmental processes [93–95].



**Figure 2.** STRING interaction enrichment analysis identified two dominant enrichment networks as denoted by red (synaptic membranes—(A)) and yellow nodes (ubiquitin pathway—(B)). Coloured nodes represent query proteins and the first shell of interactions, connected by coloured lines indicating the known interactions derived from experimental data, curated databases, co-expression studies, or text-mining. Blue lines denote known interactions from curated databases, pink lines represent experimentally determined interactions, black lines indicate co-expression, and purple lines signify homology.

### 3. Discussion

This study employed GS to identify genetic variants associated with ASD in a cohort from Oman, addressing the underrepresentation of genetic studies on ASD in the Middle East. The discovery of ASD genes has been hampered by high genetic heterogeneity, with many genes each contributing to small, complex effects that require large sample sizes to detect. ASD is also linked to rare variants in the general population and de novo variants in affected individuals, which are challenging to detect without diverse populations and advanced analytical approaches. Especially in Arab populations, genetic studies on ASD remain fragmentary compared to the extensive research conducted in Western populations, limiting insights into ASD's genetic landscape in this region. To address this gap, we performed GS on 104 unrelated families affected by ASD of unknown etiology in Oman, employing a trio analysis, which sequences the genomes of an affected individual and both unaffected biological parents. This approach offers several advantages over singleton sequencing, particularly in identifying de novo and inherited mutations. By examining inheritance patterns, we can classify variants as benign or pathogenic more effectively, which aids in pinpointing those most likely associated with ASD. Comparing the proband's genome with that of both parents also helps filter out common variants, thereby reducing the number of irrelevant findings. Finally, for families with suspected rare syndromic ASD, trio sequencing significantly improves diagnostic yield by uncovering rare or novel variants that singleton sequencing might miss.

Our research identified 116 unique genes across 95 families, while 10 CNVs were detected in 8 families. Of the 116 genes identified, 83 are considered strong candidate genes

for ASD based on ACMG guidelines, comprising 15 de novo, 55 homozygous, 6 compound heterozygous, and 14 hemizygous variants. Among these 83 genes, 48 are novel candidate genes significantly associated with ASD in the Omani cohort, while the remaining 35 are known ASD genes. Ten de novo micro-CNVs identified in eight ASD subjects are either intragenic or affect one end of a gene, involving a total of ten candidate genes (Table 3). Notably, at least several sporadic variants in these genes have been reported in patients with ASD, intellectual disability (ID), developmental disorders, or other NDDs in the literature, further supporting their potential role as ASD candidate genes.

STRING protein–protein interaction enrichment analysis of these genes provided molecular etiological insights, demonstrating connections between ASD and fundamental cellular functions such as presynaptic membrane signaling and ubiquitin-conjugating enzyme pathways.

Most of the variants identified in our limited cohort were either absent in the gnomAD and other polymorphism databases or present at very low frequencies (<0.01%). Furthermore, these variants displayed very low frequencies in the Qatar Genome Program (QGP) database, which includes genomic data from over 10,000 healthy individuals from Qatar [21]. The strength of our methodology was further validated by identifying variants in NDD genes or ASD genes, as documented in high-throughput sequencing databases and the literature.

An important feature of our Omani cohort was that, out of 104 families, 56 self-reported consanguineous marriages. Among these 56 families, we identified 53 strong candidate genes across 41 families. Of these 53, 47 variants were inherited from parents as either homozygous recessive (41 genes) or X-linked (6 genes). Interestingly, we also identified de novo variants in 6 families, indicating that consanguineous families can harbor both de novo and recessive variants [96]. A higher prevalence of recessive mutations suggests that consanguinity plays an important role in the genetic etiology of ASD in this region.

One aim of this study was to identify variants specific to the Arab/Oman population. Interestingly, we found seven genes with putative causative variants recurring across families: six genes appeared in two families each, and one gene *DNAH17* (dynein axonemal heavy chain 17, MIM 610063) was identified in three different families (family numbers 31, 66, and 84) with compound heterozygous variants. The nucleotide variants were identical in families 31 and 84 (c.2121C>A and c.12650T>C) but differed in family 66 (c.12389C>T and c.11825G>A). Loss of function mutations in *DNAH17* are linked to a genetically heterogeneous disorder leading to male infertility and multiple morphological abnormalities of the flagella (MMAF) [97]. However, mutations in the *DNAH17* potentially affect neuronal function, synaptic connectivity, or ciliary-related signaling pathways important for brain development [98,99]. This is supported by sporadic missense and nonsense variants in ASD patients documented in the HGMD [30,31,100].

An *FMR1* variant c.716C>T was identified in two families (22 and 91). While CGG repeat expansions in *FMR1* are known to cause ID, point mutations in *FMR1* have also been reported in ID subjects [101,102]. The *FMR1* gene (MIM 309550), located on the X-chromosome, is one of the most extensively studied genes in relation to ASD, primarily due to its role in Fragile X Syndrome (FXS) (MIM 300624) [103,104]. *FMR1* encodes the Fragile X Mental Retardation Protein (FMRP), a key regulator of synaptic plasticity and local protein synthesis at synapses [104]. In individuals with FXS, the absence or deficiency of FMRP impairs synaptic function, leading to ID, ASD, and other NDDs [105].

*SCN2A* (MIM 182390) encodes the alpha subunit of the Nav1.2 voltage-gated sodium channel, which is critical for neuronal function, particularly in initiating and propagating action potentials. Variants in *SCN2A* have been strongly associated with NDDs, including ASD [106–109]. In our study, two different de novo *SCN2A* variants were identified: c.3932T>G in family 49 and c.532G>T in family 90.

*GIPR* (MIM 137241) encodes the receptor for gastric inhibitory polypeptide, a hormone primarily involved in regulating insulin secretion and lipid metabolism [110,111]. *GIPR* plays a critical role in metabolic regulation, the gut–brain axis, and potential neuroinflam-

matory mechanisms [110–112]. Although *GIPR*'s primary functions are linked to metabolic pathways, disruptions in these processes may affect neurodevelopment and brain function, potentially contributing to ASD pathophysiology. We identified two *GIPR* variants: c.1152G>T and c.302G>A in family 96 and c.1264C>T in family 100.

The same variant c.1879C>T in *SPATA31A3* was identified in family 44 and family 58. *SPATA31A3*, a member of the SPATA (Spermatogenesis-Associated) gene family, is involved in multiple cellular processes, including DNA repair and the maintenance of genomic stability, both of which are crucial for normal brain development and function [113,114]. Two sporadic missense variants in this gene have been reported in ASD patients [30].

We identified an identical *TRIM15* variant (c.608A>G) in families 18 and 80 and 1 same variant in *TRIM73* (c.487C>T) in families 3 and 77. *TRIM15* and *TRIM73* (MIM 612549) are members of the TRIM (Tripartite Motif-containing) protein family, which participates in various biological processes, including immune response, cell proliferation, and gene expression regulation [115–118]. While the exact roles of *TRIM15* and *TRIM73* in ASD are not fully understood, emerging evidence suggests their potential involvement in neurodevelopment and immune regulation, both of which are relevant to ASD pathology [115–118]. A regulatory variant in *TRIM15* and a missense variant in *TRIM73* have been reported in ASD patients [30,31].

Identifying multiple variants of the same gene in independent patients strengthens the evidence for the gene's potential role in disease etiology. This recurrence suggests that the gene may be functionally important to the condition of ASD and helps distinguish disease genes from incidental findings. Recurrent variants in the same gene across different patients also guide further functional studies, improve genetic diagnosis, and may support the development of targeted therapeutic approaches. Furthermore, the identification of multiple genes linked to ASD within a specific population provides insights into the unique genetic architecture and risk factors of ASD in that population. Such findings can reveal population-specific variants and gene–environment interactions that may be absent or less prevalent in other groups. This knowledge improves the accuracy of genetic diagnoses and risk assessments within that population, facilitates tailored genetic counselling, and contributes to the development of population-specific treatments or interventions. Additionally, understanding these genetic variations enhances global ASD research by adding diversity to genetic databases, which is essential for identifying both shared and unique mechanisms underlying ASD across populations.

## 4. Materials and Methods

### 4.1. Sampling Procedures and DNA Isolation

Bio-samples collected for this study mainly consisted of blood from affected individuals and their parents. Samples were collected at Sultan Qaboos University, Muscat, Oman, in tubes treated with anti-coagulant ethylenediamine tetra acetic acid (EDTA) to preserve blood for DNA and RNA extraction. Genomic DNA was extracted from peripheral blood leukocytes using the Flexigene DNA extraction kit protocols (Qiagen, Hilden, Germany). DNA concentrations were measured using a Nanodrop Spectrophotometer 1000 (ND-1000; Thermo Fisher Scientific, Waltham, MA, USA) and further validated for quality using a Qubit fluorometer (ThermoFisher Scientific, Waltham, MA, USA). Finally, DNA samples were diluted and aliquoted to 2 µg per sample in new barcoded vials, which were labeled and prepared for shipment to Qatar Genome Program in Qatar for genome sequencing.

### 4.2. Library Construction and Genome Sequencing

GS for this study adhered to established protocols, incorporating rigorous quality control to ensure accuracy, reliability, and reproducibility. These measures reduce errors and prevent sample contamination to generate high-quality data [21]. Library construction and sequencing were conducted at the Sidra Clinical Genomics Laboratory Sequencing Facility, utilizing the Agilent SureSelectXT kit from Agilent Technologies, Santa Clara, CA, USA, following the manufacturer's instructions. The workflow began with mechanical



fragmentation of 200 ng of genomic DNA using the Covaris E220 ultrasonicator system, Covaris PerkinElmer, Waltham, MA, USA followed by DNA purification with AMPure XP magnetic beads, Beckman Coulter, Indianapolis, IN, USA. The fragmented DNA was subsequently end-repaired, adenylated, and ligated to SureSelect DNA adapters, Agilent Technologies, Santa Clara, CA, USA. After ligation, the DNA was further purified and amplified by PCR. The amplified library was then hybridized to biotin- labeled probes targeting specific regions of interest, captured with streptavidin-coated beads, Agilent Technologies, Santa Clara, CA, USA and subjected to a second round of PCR with a unique index. Library quality was assessed using the Agilent 2100 Bioanalyzer, and concentrations were quantified with the Qubit system, ThermoFisher, Waltham, MA, USA. Libraries meeting quality control criteria were pooled and sequenced on the Illumina HiSeq 4000 platform, generating a minimum of 50 million paired-end reads ( $2 \times 150$  bp) per sample.

#### 4.3. Variant Calling Process

Sidra Research, Clinical Genomics Laboratory (CGL) in Qatar provided GS data in Fastq format. This raw FastQ data were further processed using Illumina's DRAGEN (Dynamic Read Analysis for GENomics) platform, (v.4.2.4) a high-performance solution optimized for GS data analysis [59,60]. DRAGEN pipeline starts by ingesting raw sequencing data in Fastq format by applying initial quality checks to assess read quality metrics such as base quality scores, GC content, and adapter content. Subsequently, the reads are aligned to reference genome (GRCh38) using DRAGEN's hardware-accelerated mapping and alignment algorithm, known as "hash table alignment", which enables efficient and rapid reference-based mapping. DRAGEN then identifies and marks duplicate reads, and base quality scores are recalibrated to correct for systematic biases observed during sequencing process by analysing empirical error rates across various sequence contexts. Variant calling for SNVs and insertions/deletions (indels) is conducted using probabilistic models adapted from GATK's HaplotypeCaller, optimized for DRAGEN's architecture. Beyond SNVs and indels detection, DRAGEN also identifies structural variants (SVs) in chromosomes, such as CNVs, large deletions or duplications, and translocations. Haplotype phasing is performed to assign variants to maternal or paternal chromosomes, enhancing the interpretability of variant calls and inheritance patterns, particularly in compound heterozygous variants. Finally, identified variants are annotated using databases such as ClinVar, dbSNP, and other population and clinical variant resources and presented in annotated VCF files for downstream analysis [59,60]

#### 4.4. Integrating Genetic Databases to Validate ASD Candidate Genes

To validate the significance of our ASD candidate genes, we performed gene annotation using three comprehensive resources: the SFARI gene database <https://gene.sfari.org/database/human-gene/> (accessed on 1 August 2024), the Genomics England NDD/Autism gene panel <https://panelapp.genomicsengland.co.uk/panels/285/> (accessed on 1 August 2024), and HGMD <https://my.qiagen.digitalinsights.com/bbp/view/hgmd/pro/all.php> (accessed on 1 August 2024). The SFARI database catalogs genes potentially implicated in ASD, while the Genomics England NDD/Autism panel includes genes associated with various NDDs, including autism, epilepsy, ID, attention deficit hyperactivity disorder (ADHD), and other NDDs. HGMD, on the other hand, includes both causative variants linked to known human disorders and sporadic variants identified in patients with specific disorders, even if their causality has not been definitively established. This resource offers a comprehensive collection of genetic variants reported in the literature, encompassing both pathogenic mutations and variants of uncertain significance associated with inherited diseases [89]. We cross-referenced our newly identified gene list with the following resources: SFARI, Genomics England NDD/ASD, and HGMD. Moreover, we categorized each gene as "Known" or "Novel" in Table 2 based on their presence in these databases. This comparison

aligns our results with well-established databases, enhancing the relevance of our findings and supporting the potential involvement of these genes in ASD.

#### 4.5. STRING Interaction Enrichment Analysis

The STRING 3.0 analysis was utilized for our candidate genes to explore and validate protein–protein interaction (PPI) within this unique gene set. By mapping these genes to STRING, we aimed to identify functional networks and interaction clusters that could reveal biological pathways and mechanisms potentially underlying ASD in Omani population. This approach enabled us to integrate multiple interaction sources, including text mining, experimental data, curated databases, and co-expression while restricting our analysis to “homo sapiens” and applying an interaction score threshold of >0.4 to construct the PPI networks. Finally, functional clusters within these PPI networks were subsequently identified using k-means clustering. Each node represents a protein with its 3D structure displayed, while edges signify protein–protein interactions. Blue lines denote known interactions from curated databases, pink lines represent experimentally determined interactions, black lines indicate co-expression, and purple lines signify homology. A low PPI enrichment *p*-value suggests a non-random network, indicating that the observed number of edges is statistically significant.

## 5. Conclusions

This study identified ASD candidate genes in an Omani cohort, shedding light on the genetic underpinnings of ASD in this population. We have identified 48 novel ASD candidates and 35 genes previously linked to NDDs. The presence of multiple candidate variants in some individuals suggests potential digenic or polygenic inheritance. The high prevalence of homozygous variants, 61% in the entire cohort and 84% in consanguineous families supports a recessive genetic architecture for many ASD cases in this population, consistent with the high rates of consanguinity.

These findings underscore the importance of considering population-specific genetic factors and suggest that recessive alleles contribute significantly to ASD risk in this region. This work also highlights the value of examining rare, potentially pathogenic variants in consanguineous populations, which may improve genetic counseling and inform public health initiatives, including genetic screening in Oman and the Middle East. Identifying new ASD-associated genes provides insights into the disorder’s genetic mechanisms and could guide the development of diagnostic tools, such as gene panels, for early and precise diagnosis. Ultimately, these discoveries open avenues for personalized treatment strategies, potentially enabling more targeted and effective interventions for ASD. Further research on these genetic pathways will be essential for developing therapies that improve outcomes for individuals with ASD.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms252413700/s1>.

**Author Contributions:** Conceptualization, W.A.-M. and L.W.S.; data curation, V.G., A.B.-M., A.B.I., W.H., A.A., and W.A.-M.; formal analysis, V.G., A.B.-M., J.-J.H., A.A., H.-G.K., and W.A.-M.; funding acquisition, L.W.S.; investigation, V.G. and A.B.-M.; methodology, V.G., A.B.-M., J.-J.H., W.H., A.A., and H.-G.K.; project administration, A.B.I., W.A.-M., and L.W.S.; resources, A.B.I., W.A.-M., and L.W.S.; software, V.G. and A.B.-M.; supervision, W.A.-M.; validation, V.G. and A.B.-M.; writing—original draft, V.G., A.B.-M., and L.W.S.; writing—review and editing, V.G., A.B.I., J.-J.H., A.A., H.-G.K., W.A.-M., and L.W.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Qatar Foundation (Qatar Foundation for Education, Science and Community Development), Doha, Qatar-QB06.

**Institutional Review Board Statement:** This study was approved by the Qatar Biomedical Research Institute Review Board at Hamad Bin Khalifa University in Qatar for project titled—“Identification of Genetic Variants Associated with Autism Spectrum Disorder” with IRB Protocol Reference number-

QBRI-IRB-2022-7 and the Medical Research Ethics Committee (MREC) at Sultan Qaboos University in Oman for project—“Whole Genome Sequencing (WGS) for children with Autism Spectrum Disorder (ASD)” with REF. NO. SQU-EC/508/2021, MREC #2505.

**Informed Consent Statement:** Written informed consent was obtained from all participants involved in this study.

**Data Availability Statement:** Data are available upon request from corresponding authors.

**Acknowledgments:** The authors would like to acknowledge Weill Cornell Bioinformatics core facility-(Gaurav Thareja, Tanwir Habib, and Karsten Suhre) for helping with the VCF file generation for genome sequencing data.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

autism spectrum disorder (ASD), neurodevelopmental disorders (NDDs), intellectual disability (ID), next-generation sequencing (NGS), human gene mutation database (HGMD), genome sequencing (GS), single nucleotide variants (SNVs), copy number variants (CNVs), structural variants (SVs), Qatar-Genome Project (QGP), allele frequencies (AF).

## Web Resources:

OMIM, <https://www.omim.org/> (accessed on 1 August 2024).

HGMD, <https://my.qiagen.digitalinsights.com/bbp/view/hgmd/pro/all.php> (accessed on 1 August 2024).

Genomics England NDD/autism panel genes <https://panelapp.genomicsengland.co.uk/panels/285/> (accessed on 1 August 2024).

SFARI (Simons Foundation Autism Research Initiative) <https://gene.sfari.org/database/human-gene/> (accessed on 1 August 2024).

## References

1. First, M.B. Diagnostic and statistical manual of mental disorders, 5th edition, and clinical utility. *J. Nerv. Ment. Dis.* **2013**, *201*, 727–729. [[CrossRef](#)] [[PubMed](#)]
2. Mitchell, K.J. The genetics of neurodevelopmental disease. *Curr. Opin. Neurobiol.* **2011**, *21*, 197–203. [[CrossRef](#)] [[PubMed](#)]
3. Levy, S.E.; Mandell, D.S.; Schultz, R.T. Autism. *Lancet* **2009**, *374*, 1627–1638. [[CrossRef](#)] [[PubMed](#)]
4. Malachowski, M. Understanding Mental Disorders: Your Guide to DSM-5, by the American Psychiatric Association. *Med. Ref. Serv. Q.* **2016**, *35*, 467–468. [[CrossRef](#)] [[PubMed](#)]
5. Hyman, S.L.; Levy, S.E.; Myers, S.M. Council on Children with Disabilities, S.O.D.; Behavioral, P. Identification, Evaluation, and Management of Children with Autism Spectrum Disorder. *Pediatrics* **2020**, *145*, e20193447. [[CrossRef](#)]
6. Devnani, P.A.; Hegde, A.U. Autism and sleep disorders. *J. Pediatr. Neurosci.* **2015**, *10*, 304–307. [[CrossRef](#)] [[PubMed](#)]
7. Besag, F.M. Epilepsy in patients with autism: Links, risks and treatment challenges. *Neuropsychiatr. Dis. Treat.* **2018**, *14*, 1–10. [[CrossRef](#)] [[PubMed](#)]
8. Bjorklund, G.; Pivina, L.; Dadar, M.; Meguid, N.A.; Semenova, Y.; Anwar, M.; Chirumbolo, S. Gastrointestinal alterations in autism spectrum disorder: What do we know? *Neurosci. Biobehav. Rev.* **2020**, *118*, 111–120. [[CrossRef](#)]
9. Junaid, M.; Slack-Smith, L.; Wong, K.; Bourke, J.; Baynam, G.; Calache, H.; Leonard, H. Association between craniofacial anomalies, intellectual disability and autism spectrum disorder: Western Australian population-based study. *Pediatr. Res.* **2022**, *92*, 1795–1804. [[CrossRef](#)] [[PubMed](#)]
10. Razali, R.M.; Rodriguez-Flores, J.; Ghorbani, M.; Naeem, H.; Aamer, W.; Aliyev, E.; Jubran, A.; Qatar Genome Program Research, C.; Clark, A.G.; Fakhro, K.A.; et al. Thousands of Qatari genomes inform human migration history and improve imputation of Arab haplotypes. *Nat. Commun.* **2021**, *12*, 5929. [[CrossRef](#)] [[PubMed](#)]
11. Ben-Omran, T.; Al Ghanim, K.; Yavarna, T.; El Akoum, M.; Samara, M.; Chandra, P.; Al-Dewik, N. Effects of consanguinity in a cohort of subjects with certain genetic disorders in Qatar. *Mol. Genet. Genom. Med.* **2020**, *8*, e1051. [[CrossRef](#)]
12. El Mouzan, M.I.; Al Salloum, A.A.; Al Herbish, A.S.; Qurachi, M.M.; Al Omar, A.A. Consanguinity and major genetic disorders in Saudi children: A community-based cross-sectional study. *Ann. Saudi Med.* **2008**, *28*, 169–173. [[CrossRef](#)] [[PubMed](#)]
13. Hamamy, H.; Antonarakis, S.E.; Cavalli-Sforza, L.L.; Temtamy, S.; Romeo, G.; Kate, L.P.; Bennett, R.L.; Shaw, A.; Megarbane, A.; van Duijn, C.; et al. Consanguineous marriages, pearls and perils: Geneva International Consanguinity Workshop Report. *Genet. Med.* **2011**, *13*, 841–847. [[CrossRef](#)] [[PubMed](#)]
14. Hamamy, H.A.; Masri, A.T.; Al-Hadidy, A.M.; Ajlouni, K.M. Consanguinity and genetic disorders. Profile from Jordan. *Saudi Med. J.* **2007**, *28*, 1015–1017. [[PubMed](#)]

15. Bizzari, S.; Nair, P.; Hana, S.; Deepthi, A.; Al-Ali, M.T.; Al-Gazali, L.; El-Hayek, S. Spectrum of genetic disorders and gene variants in the United Arab Emirates national population: Insights from the CTGA database. *Front. Genet.* **2023**, *14*, 1177204. [[CrossRef](#)] [[PubMed](#)]
16. Khayat, A.M.; Alshareef, B.G.; Alharbi, S.F.; AlZahrani, M.M.; Alshangity, B.A.; Tashkandi, N.F. Consanguineous Marriage and Its Association with Genetic Disorders in Saudi Arabia: A Review. *Cureus* **2024**, *16*, e53888. [[CrossRef](#)]
17. Kilshaw, S.; Al Raisi, T.; Alshaban, F. Arranging marriage; negotiating risk: Genetics and society in Qatar. *Anthropol. Med.* **2015**, *22*, 98–113. [[CrossRef](#)] [[PubMed](#)]
18. El Naofal, M.; Ramaswamy, S.; Alsarhan, A.; Nugud, A.; Sarfraz, F.; Janbaz, H.; Taylor, A.; Jain, R.; Halabi, N.; Yaslam, S.; et al. The genomic landscape of rare disorders in the Middle East. *Genome Med.* **2023**, *15*, 5. [[CrossRef](#)] [[PubMed](#)]
19. Fatumo, S.; Chikowore, T.; Choudhury, A.; Ayub, M.; Martin, A.R.; Kuchenbaecker, K. A roadmap to increase diversity in genomic studies. *Nat. Med.* **2022**, *28*, 243–250. [[CrossRef](#)] [[PubMed](#)]
20. Lemke, A.A.; Esplin, E.D.; Goldenberg, A.J.; Gonzaga-Jauregui, C.; Hanchard, N.A.; Harris-Wai, J.; Ideozu, J.E.; Isasi, R.; Landstrom, A.P.; Prince, A.E.R.; et al. Addressing underrepresentation in genomics research through community engagement. *Am. J. Hum. Genet.* **2022**, *109*, 1563–1571. [[CrossRef](#)] [[PubMed](#)]
21. Mbarek, H.; Devadoss Gandhi, G.; Selvaraj, S.; Al-Muftah, W.; Badji, R.; Al-Sarraj, Y.; Saad, C.; Darwish, D.; Alvi, M.; Fadl, T.; et al. Qatar genome: Insights on genomics from the Middle East. *Hum. Mutat.* **2022**, *43*, 499–510. [[CrossRef](#)] [[PubMed](#)]
22. Wojcik, G.L.; Graff, M.; Nishimura, K.K.; Tao, R.; Haessler, J.; Gignoux, C.R.; Highland, H.M.; Patel, Y.M.; Sorokin, E.P.; Avery, C.L.; et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **2019**, *570*, 514–518. [[CrossRef](#)] [[PubMed](#)]
23. Alshaban, F.; Aldosari, M.; Al-Shammari, H.; El-Hag, S.; Ghazal, I.; Tolefat, M.; Ali, M.; Kamal, M.; Abdel Aati, N.; Abeidah, M.; et al. Prevalence and correlates of autism spectrum disorder in Qatar: A national study. *J. Child Psychol. Psychiatry* **2019**, *60*, 1254–1268. [[CrossRef](#)]
24. Abdi, M.; Aliyev, E.; Trost, B.; Kohailan, M.; Aamer, W.; Syed, N.; Shaath, R.; Gandhi, G.D.; Engchuan, W.; Howe, J.; et al. Genomic architecture of autism spectrum disorder in Qatar: The BARAKA-Qatar Study. *Genome Med.* **2023**, *15*, 81. [[CrossRef](#)] [[PubMed](#)]
25. Al-Sarraj, Y.; Taha, R.Z.; Al-Dous, E.; Ahram, D.; Abbasi, S.; Abuazab, E.; Shaath, H.; Habbab, W.; Errafii, K.; Bejaoui, Y.; et al. The genetic landscape of autism spectrum disorder in the Middle Eastern population. *Front. Genet.* **2024**, *15*, 1363849. [[CrossRef](#)] [[PubMed](#)]
26. Ben-Mahmoud, A.; Gupta, V.; Abdelaleem, A.; Thompson, R.; Aden, A.; Mbarek, H.; Saad, C.; Tolefat, M.; Alshaban, F.; Stanton, L.W.; et al. Genome Sequencing Identifies 13 Novel Candidate Risk Genes for Autism Spectrum Disorder in a Qatari Cohort. *Int. J. Mol. Sci.* **2024**, *25*, 11551. [[CrossRef](#)] [[PubMed](#)]
27. Gupta, V.; Ben-Mahmoud, A.; Ku, B.; Velayutham, D.; Jan, Z.; Yousef Aden, A.; Kubbar, A.; Alshaban, F.; Stanton, L.W.; Jithesh, P.V.; et al. Identification of two novel autism genes, TRPC4 and SCFD2, in Qatar simplex families through exome sequencing. *Front. Psychiatry* **2023**, *14*, 1251884. [[CrossRef](#)] [[PubMed](#)]
28. Al-Mamri, W.; Idris, A.B.; Dakak, S.; Al-Shekaili, M.; Al-Harathi, Z.; Alnaamani, A.M.; Alhinai, F.I.; Jalees, S.; Al Hatmi, M.; El-Naggari, M.A.; et al. Revisiting the Prevalence of Autism Spectrum Disorder among Omani Children: A multicentre study. *Sultan Qaboos Univ. Med. J.* **2019**, *19*, e305–e309. [[CrossRef](#)]
29. Satterstrom, F.K.; Kosmicki, J.A.; Wang, J.; Breen, M.S.; De Rubeis, S.; An, J.Y.; Peng, M.; Collins, R.; Grove, J.; Klei, L.; et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* **2020**, *180*, 568–584 e523. [[CrossRef](#)] [[PubMed](#)]
30. Fu, J.M.; Satterstrom, F.K.; Peng, M.; Brand, H.; Collins, R.L.; Dong, S.; Wamsley, B.; Klei, L.; Wang, L.; Hao, S.P.; et al. Rare coding variation provides insight into the genetic architecture and phenotypic context of autism. *Nat. Genet.* **2022**, *54*, 1320–1331. [[CrossRef](#)] [[PubMed](#)]
31. Zhou, X.; Feliciano, P.; Shu, C.; Wang, T.; Astrovskaya, I.; Hall, J.B.; Obiajulu, J.U.; Wright, J.R.; Murali, S.C.; Xu, S.X.; et al. Integrating de novo and inherited variants in 42,607 autism cases identifies mutations in new moderate-risk genes. *Nat. Genet.* **2022**, *54*, 1305–1319. [[CrossRef](#)] [[PubMed](#)]
32. De Rubeis, S.; He, X.; Goldberg, A.P.; Poultney, C.S.; Samocha, K.; Cicek, A.E.; Kou, Y.; Liu, L.; Fromer, M.; Walker, S.; et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **2014**, *515*, 209–215. [[CrossRef](#)] [[PubMed](#)]
33. Huguët, G.; Ey, E.; Bourgeron, T. The genetic landscapes of autism spectrum disorders. *Annu. Rev. Genom. Hum. Genet.* **2013**, *14*, 191–213. [[CrossRef](#)] [[PubMed](#)]
34. Postorino, V.; Fatta, L.M.; Sanges, V.; Giovagnoli, G.; De Peppo, L.; Vicari, S.; Mazzone, L. Intellectual disability in Autism Spectrum Disorder: Investigation of prevalence in an Italian sample of children and adolescents. *Res. Dev. Disabil.* **2016**, *48*, 193–201. [[CrossRef](#)] [[PubMed](#)]
35. Sanders, S.J.; He, X.; Willsey, A.J.; Ercan-Sencicek, A.G.; Samocha, K.E.; Cicek, A.E.; Murtha, M.T.; Bal, V.H.; Bishop, S.L.; Dong, S.; et al. Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron* **2015**, *87*, 1215–1233. [[CrossRef](#)] [[PubMed](#)]
36. Sandin, S.; Lichtenstein, P.; Kuja-Halkola, R.; Hultman, C.; Larsson, H.; Reichenberg, A. The Heritability of Autism Spectrum Disorder. *JAMA* **2017**, *318*, 1182–1184. [[CrossRef](#)] [[PubMed](#)]

37. Hnoonual, A.; Thammachote, W.; Tim-Aroon, T.; Rojnueangnit, K.; Hansakunachai, T.; Sombuntham, T.; Roongpraiwan, R.; Worachotekamjorn, J.; Chuthapisith, J.; Fucharoen, S.; et al. Chromosomal microarray analysis in a cohort of underrepresented population identifies SERINC2 as a novel candidate gene for autism spectrum disorder. *Sci. Rep.* **2017**, *7*, 12096. [[CrossRef](#)] [[PubMed](#)]
38. Turner, T.N.; Wilfert, A.B.; Bakken, T.E.; Bernier, R.A.; Pepper, M.R.; Zhang, Z.; Torene, R.I.; Retterer, K.; Eichler, E.E. Sex-Based Analysis of De Novo Variants in Neurodevelopmental Disorders. *Am. J. Hum. Genet.* **2019**, *105*, 1274–1285. [[CrossRef](#)]
39. Grove, J.; Ripke, S.; Als, T.D.; Mattheisen, M.; Walters, R.K.; Won, H.; Pallesen, J.; Agerbo, E.; Andreassen, O.A.; Anney, R.; et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **2019**, *51*, 431–444. [[CrossRef](#)]
40. Guo, H.; Duyzend, M.H.; Coe, B.P.; Baker, C.; Hoekzema, K.; Gerds, J.; Turner, T.N.; Zody, M.C.; Beighley, J.S.; Murali, S.C.; et al. Genome sequencing identifies multiple deleterious variants in autism patients with more severe phenotypes. *Genet. Med.* **2019**, *21*, 1611–1620. [[CrossRef](#)]
41. Guo, H.; Wang, T.; Wu, H.; Long, M.; Coe, B.P.; Li, H.; Xun, G.; Ou, J.; Chen, B.; Duan, G.; et al. Inherited and multiple de novo mutations in autism/developmental delay risk genes suggest a multifactorial model. *Mol. Autism* **2018**, *9*, 64. [[CrossRef](#)]
42. Hogg, G.; Severson, E.A.; Cai, L.; Hoffmann, H.M.; Holden, K.A.; Fitzgerald, K.; Kenyon, A.; Zeng, Q.; Mooney, M.; Gardner, S.; et al. Clinical characterization of the mutational landscape of 24,639 real-world samples from patients with myeloid malignancies. *Cancer Genet.* **2023**, 278–279, 38–49. [[CrossRef](#)] [[PubMed](#)]
43. Iossifov, I.; O’Roak, B.J.; Sanders, S.J.; Ronemus, M.; Krumm, N.; Levy, D.; Stessman, H.A.; Witherspoon, K.T.; Vives, L.; Patterson, K.E.; et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **2014**, *515*, 216–221. [[CrossRef](#)]
44. Nishioka, M.; Kazuno, A.A.; Nakamura, T.; Sakai, N.; Hayama, T.; Fujii, K.; Matsuo, K.; Komori, A.; Ishiwata, M.; Watanabe, Y.; et al. Systematic analysis of exonic germline and postzygotic de novo mutations in bipolar disorder. *Nat. Commun.* **2021**, *12*, 3750. [[CrossRef](#)] [[PubMed](#)]
45. O’Roak, B.J.; Deriziotis, P.; Lee, C.; Vives, L.; Schwartz, J.J.; Girirajan, S.; Karakoc, E.; Mackenzie, A.P.; Ng, S.B.; Baker, C.; et al. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat. Genet.* **2011**, *43*, 585–589. [[CrossRef](#)]
46. O’Roak, B.J.; Stessman, H.A.; Boyle, E.A.; Witherspoon, K.T.; Martin, B.; Lee, C.; Vives, L.; Baker, C.; Hiatt, J.B.; Nickerson, D.A.; et al. Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nat. Commun.* **2014**, *5*, 5595. [[CrossRef](#)] [[PubMed](#)]
47. Sanders, S.J.; Murtha, M.T.; Gupta, A.R.; Murdoch, J.D.; Raubeson, M.J.; Willsey, A.J.; Ercan-Sencicek, A.G.; DiLullo, N.M.; Parikshak, N.N.; Stein, J.L.; et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **2012**, *485*, 237–241. [[CrossRef](#)]
48. Takata, A.; Miyake, N.; Tsurusaki, Y.; Fukai, R.; Miyatake, S.; Koshimizu, E.; Kushima, I.; Okada, T.; Morikawa, M.; Uno, Y.; et al. Integrative Analyses of De Novo Mutations Provide Deeper Biological Insights into Autism Spectrum Disorder. *Cell Rep.* **2018**, *22*, 734–747. [[CrossRef](#)]
49. Takata, A.; Nakashima, M.; Saitsu, H.; Mizuguchi, T.; Mitsuhashi, S.; Takahashi, Y.; Okamoto, N.; Osaka, H.; Nakamura, K.; Tohyama, J.; et al. Comprehensive analysis of coding variants highlights genetic complexity in developmental and epileptic encephalopathy. *Nat. Commun.* **2019**, *10*, 2506. [[CrossRef](#)]
50. Valentino, F.; Bruno, L.P.; Doddato, G.; Giliberti, A.; Tita, R.; Resciniti, S.; Fallerini, C.; Bruttini, M.; Lo Rizzo, C.; Mencarelli, M.A.; et al. Exome Sequencing in 200 Intellectual Disability/Autistic Patients: New Candidates and Atypical Presentations. *Brain Sci.* **2021**, *11*, 936. [[CrossRef](#)]
51. Warriar, V.; Zhang, X.; Reed, P.; Havdahl, A.; Moore, T.M.; Cliquet, F.; Leblond, C.S.; Rolland, T.; Rosengren, A.; Eu-Aims, L.; et al. Genetic correlates of phenotypic heterogeneity in autism. *Nat. Genet.* **2022**, *54*, 1293–1304. [[CrossRef](#)]
52. Wilfert, A.B.; Turner, T.N.; Murali, S.C.; Hsieh, P.; Sulovari, A.; Wang, T.; Coe, B.P.; Guo, H.; Hoekzema, K.; Bakken, T.E.; et al. Recent ultra-rare inherited variants implicate new autism candidate risk genes. *Nat. Genet.* **2021**, *53*, 1125–1134. [[CrossRef](#)]
53. Trost, B.; Thiruvahindrapuram, B.; Chan, A.J.S.; Engchuan, W.; Higginbotham, E.J.; Howe, J.L.; Loureiro, L.O.; Reuter, M.S.; Roshandel, D.; Whitney, J.; et al. Genomic architecture of autism from comprehensive whole-genome sequence annotation. *Cell* **2022**, *185*, 4409–4427.e4418. [[CrossRef](#)]
54. Gaugler, T.; Klei, L.; Sanders, S.J.; Bodea, C.A.; Goldberg, A.P.; Lee, A.B.; Mahajan, M.; Manaa, D.; Pawitan, Y.; Reichert, J.; et al. Most genetic risk for autism resides with common variation. *Nat. Genet.* **2014**, *46*, 881–885. [[CrossRef](#)]
55. Lim, E.T.; Raychaudhuri, S.; Sanders, S.J.; Stevens, C.; Sabo, A.; MacArthur, D.G.; Neale, B.M.; Kirby, A.; Ruderfer, D.M.; Fromer, M.; et al. Rare complete knockouts in humans: Population distribution and significant role in autism spectrum disorders. *Neuron* **2013**, *77*, 235–242. [[CrossRef](#)]
56. Tuncay, I.O.; Parmalee, N.L.; Khalil, R.; Kaur, K.; Kumar, A.; Jimale, M.; Howe, J.L.; Goodspeed, K.; Evans, P.; Alzghoul, L.; et al. Analysis of recent shared ancestry in a familial cohort identifies coding and noncoding autism spectrum disorder variants. *NPJ Genom. Med.* **2022**, *7*, 13. [[CrossRef](#)]
57. Choi, L.; An, J.Y. Genetic architecture of autism spectrum disorder: Lessons from large-scale genomic studies. *Neurosci. Biobehav. Rev.* **2021**, *128*, 244–257. [[CrossRef](#)] [[PubMed](#)]
58. Doan, R.N.; Lim, E.T.; De Rubeis, S.; Betancur, C.; Cutler, D.J.; Chiocchetti, A.G.; Overman, L.M.; Soucy, A.; Goetze, S.; Autism Sequencing, C.; et al. Recessive gene disruptions in autism spectrum disorder. *Nat. Genet.* **2019**, *51*, 1092–1098. [[CrossRef](#)] [[PubMed](#)]

59. Behera, S.; Catreux, S.; Rossi, M.; Truong, S.; Huang, Z.; Ruehle, M.; Visvanath, A.; Parnaby, G.; Roddey, C.; Onuchic, V.; et al. Comprehensive genome analysis and variant detection at scale using DRAGEN. *Nat. Biotechnol.* **2024**, 1–15. [[CrossRef](#)] [[PubMed](#)]
60. Olson, N.D.; Wagner, J.; McDaniel, J.; Stephens, S.H.; Westreich, S.T.; Prasanna, A.G.; Johanson, E.; Boja, E.; Maier, E.J.; Serang, O.; et al. PrecisionFDA Truth Challenge V2: Calling variants from short and long reads in difficult-to-map regions. *Cell. Genom.* **2022**, 2, 1–12. [[CrossRef](#)]
61. Kostic, M.; Raymond, J.J.; Freyre, C.A.C.; Henry, B.; Tumkaya, T.; Khlgatyan, J.; Dvornik, J.; Li, J.; Hsiao, J.S.; Cheon, S.H.; et al. Patient Brain Organoids Identify a Link between the 16p11.2 Copy Number Variant and the RBFOX1 Gene. *ACS Chem. Neurosci.* **2023**, 14, 3993–4012. [[CrossRef](#)] [[PubMed](#)]
62. O’Leary, A.; Fernandez-Castillo, N.; Gan, G.; Yang, Y.; Yotova, A.Y.; Kranz, T.M.; Grunewald, L.; Freudenberg, F.; Anton-Galindo, E.; Cabana-Dominguez, J.; et al. Behavioural and functional evidence revealing the role of RBFOX1 variation in multiple psychiatric disorders and traits. *Mol. Psychiatry* **2022**, 27, 4464–4473. [[CrossRef](#)] [[PubMed](#)]
63. Fernandez-Castillo, N.; Gan, G.; van Donkelaar, M.M.J.; Vaht, M.; Weber, H.; Retz, W.; Meyer-Lindenberg, A.; Franke, B.; Harro, J.; Reif, A.; et al. RBFOX1, encoding a splicing regulator, is a candidate gene for aggressive behavior. *Eur. Neuropsychopharmacol.* **2020**, 30, 44–55. [[CrossRef](#)]
64. Hamada, N.; Ito, H.; Nishijo, T.; Iwamoto, I.; Morishita, R.; Tabata, H.; Momiyama, T.; Nagata, K. Essential role of the nuclear isoform of RBFOX1, a candidate gene for autism spectrum disorders, in the brain development. *Sci. Rep.* **2016**, 6, 30805. [[CrossRef](#)] [[PubMed](#)]
65. Altintas, M.; Yildirim, M.; Bektas, O.; Teber, S. Progressive Myoclonus Epilepsy and Beyond: Systematic Review of SEMA6B-Related Disorders. *Neuropediatrics* **2024**. [[CrossRef](#)]
66. Seiffert, S.; Pendziwiat, M.; Bierhals, T.; Goel, H.; Schwarz, N.; van der Ven, A.; Bosselmann, C.M.; Lemke, J.; Syrbe, S.; Willemsen, M.H.; et al. Modulating effects of FGF12 variants on Na(V)1.2 and Na(V)1.6 being associated with developmental and epileptic encephalopathy and Autism spectrum disorder: A case series. *EBioMedicine* **2022**, 83, 104234. [[CrossRef](#)] [[PubMed](#)]
67. Hamanaka, K.; Imagawa, E.; Koshimizu, E.; Miyatake, S.; Tohyama, J.; Yamagata, T.; Miyauchi, A.; Ekhilevitch, N.; Nakamura, F.; Kawashima, T.; et al. De Novo Truncating Variants in the Last Exon of SEMA6B Cause Progressive Myoclonic Epilepsy. *Am. J. Hum. Genet.* **2020**, 106, 549–558. [[CrossRef](#)]
68. Torene, R.I.; Guillen Sacoto, M.J.; Millan, F.; Zhang, Z.; McGee, S.; Oetjens, M.; Heise, E.; Chong, K.; Sidlow, R.; O’Grady, L.; et al. Systematic analysis of variants escaping nonsense-mediated decay uncovers candidate Mendelian diseases. *Am. J. Hum. Genet.* **2024**, 111, 70–81. [[CrossRef](#)] [[PubMed](#)]
69. Cheng, Y.Y.; Chang, K.C.; Chen, P.L.; Yeung, C.Y.; Liou, B.Y.; Chen, H.L. SLCO1B1 and SLCO1B3 genetic mutations in Taiwanese patients with Rotor syndrome. *J. Formos. Med. Assoc.* **2023**, 122, 648–652. [[CrossRef](#)] [[PubMed](#)]
70. Jakobs, P.M.; Hess, J.F.; FitzGerald, P.G.; Kramer, P.; Weleber, R.G.; Litt, M. Autosomal-dominant congenital cataract associated with a deletion mutation in the human beaded filament protein gene BFSP2. *Am. J. Hum. Genet.* **2000**, 66, 1432–1436. [[CrossRef](#)]
71. Mori, D.; Inami, C.; Ikeda, R.; Sawahata, M.; Urata, S.; Yamaguchi, S.T.; Kobayashi, Y.; Fujita, K.; Arioka, Y.; Okumura, H.; et al. Mice with deficiency in Pcdh15, a gene associated with bipolar disorders, exhibit significantly elevated diurnal amplitudes of locomotion and body temperature. *Transl. Psychiatry* **2024**, 14, 216. [[CrossRef](#)] [[PubMed](#)]
72. Nardecchia, F.; De Giorgi, A.; Palombo, F.; Fiorini, C.; De Negri, A.M.; Carelli, V.; Caporali, L.; Leuzzi, V. Missense PDSS1 mutations in CoenzymeQ10 synthesis cause optic atrophy and sensorineural deafness. *Ann. Clin. Transl. Neurol.* **2021**, 8, 247–251. [[CrossRef](#)] [[PubMed](#)]
73. Ramos, P.S.; Sajuthi, S.; Langefeld, C.D.; Walker, S.J. Immune function genes CD99L2, JARID2 and TPO show association with autism spectrum disorder. *Mol. Autism* **2012**, 3, 4. [[CrossRef](#)] [[PubMed](#)]
74. Roelens, R.; Peigneur, A.N.F.; Voets, T.; Vriens, J. Neurodevelopmental disorders caused by variants in TRPM3. *Biochim. Biophys. Acta Mol. Cell Res.* **2024**, 1871, 119709. [[CrossRef](#)]
75. Lima Caldeira, G.; Peca, J.; Carvalho, A.L. New insights on synaptic dysfunction in neuropsychiatric disorders. *Curr. Opin. Neurobiol.* **2019**, 57, 62–70. [[CrossRef](#)]
76. Morris-Rosendahl, D.J.; Crocq, M.A. Neurodevelopmental disorders—the history and future of a diagnostic concept. *Dialogues Clin. Neurosci.* **2020**, 22, 65–72. [[CrossRef](#)] [[PubMed](#)]
77. Talkowski, M.E.; Sanders, S. Diverse mutations in autism-related genes and their expression in the developing brain. *Nat. Genet.* **2022**, 54, 1263–1264. [[CrossRef](#)]
78. Antaki, D.; Guevara, J.; Maihofer, A.X.; Klein, M.; Gujral, M.; Grove, J.; Carey, C.E.; Hong, O.; Arranz, M.J.; Hervas, A.; et al. A phenotypic spectrum of autism is attributable to the combined effects of rare variants, polygenic risk and sex. *Nat. Genet.* **2022**, 54, 1284–1292. [[CrossRef](#)]
79. Aoki, Y.; Niihori, T.; Kawame, H.; Kurosawa, K.; Ohashi, H.; Tanaka, Y.; Filocamo, M.; Kato, K.; Suzuki, Y.; Kure, S.; et al. Germline mutations in HRAS proto-oncogene cause Costello syndrome. *Nat. Genet.* **2005**, 37, 1038–1040. [[CrossRef](#)]
80. Bourgeron, T. From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nat. Rev. Neurosci.* **2015**, 16, 551–563. [[CrossRef](#)]
81. Brandler, W.M.; Antaki, D.; Gujral, M.; Kleiber, M.L.; Whitney, J.; Maile, M.S.; Hong, O.; Chapman, T.R.; Tan, S.; Tandon, P.; et al. Paternally inherited cis-regulatory structural variants are associated with autism. *Science* **2018**, 360, 327–331. [[CrossRef](#)] [[PubMed](#)]

82. Castel, S.E.; Cervera, A.; Mohammadi, P.; Aguet, F.; Reverter, F.; Wolman, A.; Guigo, R.; Iossifov, I.; Vasileva, A.; Lappalainen, T. Modified penetrance of coding variants by cis-regulatory variation contributes to disease risk. *Nat. Genet.* **2018**, *50*, 1327–1334. [[CrossRef](#)] [[PubMed](#)]
83. Chen, S.; Fragoza, R.; Klei, L.; Liu, Y.; Wang, J.; Roeder, K.; Devlin, B.; Yu, H. An interactome perturbation framework prioritizes damaging missense mutations for developmental disorders. *Nat. Genet.* **2018**, *50*, 1032–1040. [[CrossRef](#)]
84. Coe, B.P.; Stessman, H.A.F.; Sulovari, A.; Geisheker, M.R.; Bakken, T.E.; Lake, A.M.; Dougherty, J.D.; Lein, E.S.; Hormozdiari, F.; Bernier, R.A.; et al. Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. *Nat. Genet.* **2019**, *51*, 106–116. [[CrossRef](#)] [[PubMed](#)]
85. Cummings, B.B.; Karczewski, K.J.; Kosmicki, J.A.; Seaby, E.G.; Watts, N.A.; Singer-Berk, M.; Mudge, J.M.; Karjalainen, J.; Satterstrom, F.K.; O'Donnell-Luria, A.H.; et al. Transcript expression-aware annotation improves rare variant interpretation. *Nature* **2020**, *581*, 452–458. [[CrossRef](#)] [[PubMed](#)]
86. Faial, T. Perturbing autism risk genes. *Nat. Genet.* **2021**, *53*, 127. [[CrossRef](#)]
87. Neale, B.M.; Kou, Y.; Liu, L.; Ma'ayan, A.; Samocha, K.E.; Sabo, A.; Lin, C.F.; Stevens, C.; Wang, L.S.; Makarov, V.; et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* **2012**, *485*, 242–245. [[CrossRef](#)]
88. O'Roak, B.J.; Vives, L.; Girirajan, S.; Karakoc, E.; Krumm, N.; Coe, B.P.; Levy, R.; Ko, A.; Lee, C.; Smith, J.D.; et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **2012**, *485*, 246–250. [[CrossRef](#)]
89. Stenson, P.D.; Mort, M.; Ball, E.V.; Chapman, M.; Evans, K.; Azevedo, L.; Hayden, M.; Heywood, S.; Millar, D.S.; Phillips, A.D.; et al. The Human Gene Mutation Database (HGMD((R))): Optimizing its use in a clinical diagnostic or research setting. *Hum. Genet.* **2020**, *139*, 1197–1207. [[CrossRef](#)]
90. Yeo, X.Y.; Lim, Y.T.; Chae, W.R.; Park, C.; Park, H.; Jung, S. Alterations of presynaptic proteins in autism spectrum disorder. *Front. Mol. Neurosci.* **2022**, *15*, 1062878. [[CrossRef](#)] [[PubMed](#)]
91. Giovedi, S.; Corradi, A.; Fassio, A.; Benfenati, F. Involvement of synaptic genes in the pathogenesis of autism spectrum disorders: The case of synapsins. *Front. Pediatr.* **2014**, *2*, 94. [[CrossRef](#)]
92. Trobiani, L.; Meringolo, M.; Diamanti, T.; Bourne, Y.; Marchot, P.; Martella, G.; Dini, L.; Pisani, A.; De Jaco, A.; Bonsi, P. The neurologins and the synaptic pathway in Autism Spectrum Disorder. *Neurosci. Biobehav. Rev.* **2020**, *119*, 37–51. [[CrossRef](#)] [[PubMed](#)]
93. George, A.J.; Hoffiz, Y.C.; Charles, A.J.; Zhu, Y.; Mabb, A.M. A Comprehensive Atlas of E3 Ubiquitin Ligase Mutations in Neurological Disorders. *Front. Genet.* **2018**, *9*, 29. [[CrossRef](#)] [[PubMed](#)]
94. Kasherman, M.A.; Premarathne, S.; Burne, T.H.J.; Wood, S.A.; Piper, M. The Ubiquitin System: A Regulatory Hub for Intellectual Disability and Autism Spectrum Disorder. *Mol. Neurobiol.* **2020**, *57*, 2179–2193. [[CrossRef](#)] [[PubMed](#)]
95. Louros, S.R.; Osterweil, E.K. Perturbed proteostasis in autism spectrum disorders. *J. Neurochem.* **2016**, *139*, 1081–1092. [[CrossRef](#)]
96. Ghasemi, M.R.; Sadeghi, H.; Hashemi-Gorji, F.; Mirfakhraie, R.; Gupta, V.; Ben-Mahmoud, A.; Bagheri, S.; Razjouyan, K.; Salehpour, S.; Tonekaboni, S.H.; et al. Exome sequencing reveals neurodevelopmental genes in simplex consanguineous Iranian families with syndromic autism. *BMC Med. Genom.* **2024**, *17*, 196. [[CrossRef](#)] [[PubMed](#)]
97. Zhou, Y.; Yu, S.; Zhang, W. The Molecular Basis of Multiple Morphological Abnormalities of Sperm Flagella and Its Impact on Clinical Practice. *Genes* **2024**, *15*, 1315. [[CrossRef](#)] [[PubMed](#)]
98. Whitfield, M.; Thomas, L.; Bequignon, E.; Schmitt, A.; Stouvenel, L.; Montantin, G.; Tissier, S.; Duquesnoy, P.; Copin, B.; Chantot, S.; et al. Mutations in DNAH17, Encoding a Sperm-Specific Axonemal Outer Dynein Arm Heavy Chain, Cause Isolated Male Infertility Due to Asthenozoospermia. *Am. J. Hum. Genet.* **2019**, *105*, 198–212. [[CrossRef](#)] [[PubMed](#)]
99. Yu, X.; Yuan, L.; Deng, S.; Xia, H.; Tu, X.; Deng, X.; Huang, X.; Cao, X.; Deng, H. Identification of DNAH17 Variants in Han-Chinese Patients with Left-Right Asymmetry Disorders. *Front. Genet.* **2022**, *13*, 862292. [[CrossRef](#)] [[PubMed](#)]
100. Lim, E.T.; Uddin, M.; De Rubeis, S.; Chan, Y.; Kamumbu, A.S.; Zhang, X.; D'Gama, A.M.; Kim, S.N.; Hill, R.S.; Goldberg, A.P.; et al. Rates, distribution and implications of postzygotic mosaic mutations in autism spectrum disorder. *Nat. Neurosci.* **2017**, *20*, 1217–1224. [[CrossRef](#)] [[PubMed](#)]
101. Nagarathinam, I.; Chong, S.S.; Thelma, B.K.; Justin Margret, J.; Venkataraman, V.; Natarajan Padmavathy, K.; Srisailapathy, C.R.S. FMR1 gene CGG repeat distribution among the three individual cohorts with intellectual disability, autism, and primary ovarian insufficiency from Tamil Nadu, Southern India. *Adv. Genet.* **2021**, *2*, e10048. [[CrossRef](#)] [[PubMed](#)]
102. Handt, M.; Epplen, A.; Hoffjan, S.; Mese, K.; Epplen, J.T.; Dekomien, G. Point mutation frequency in the FMR1 gene as revealed by fragile X syndrome screening. *Mol. Cell Probes* **2014**, *28*, 279–283. [[CrossRef](#)] [[PubMed](#)]
103. Santa Maria, L.; Aliaga, S.; Faundes, V.; Morales, P.; Pugin, A.; Curotto, B.; Soto, P.; Pena, M.I.; Salas, I.; Alliende, M.A. FMR1 gene mutations in patients with fragile X syndrome and obligate carriers: 30 years of experience in Chile. *Genet. Res.* **2016**, *98*, e11. [[CrossRef](#)] [[PubMed](#)]
104. Sidorov, M.S.; Auerbach, B.D.; Bear, M.F. Fragile X mental retardation protein and synaptic plasticity. *Mol. Brain* **2013**, *6*, 15. [[CrossRef](#)]
105. Mila, M.; Alvarez-Mora, M.I.; Madrigal, I.; Rodriguez-Revenga, L. Fragile X syndrome: An overview and update of the FMR1 gene. *Clin. Genet.* **2018**, *93*, 197–205. [[CrossRef](#)]
106. Bae, H.G.; Wu, W.C.; Nip, K.; Gould, E.; Kim, J.H. Scn2a deletion disrupts oligodendroglia function: Implication for myelination, neural circuitry, and auditory hypersensitivity in ASD. *bioRxiv* **2024**. [[CrossRef](#)]

107. Brown, C.O.; Uy, J.A.; Murtaza, N.; Rosa, E.; Alfonso, A.; Dave, B.M.; Kilpatrick, S.; Cheng, A.A.; White, S.H.; Scherer, S.W.; et al. Disruption of the autism-associated gene SCN2A alters synaptic development and neuronal signaling in patient iPSC-glutamatergic neurons. *Front. Cell Neurosci.* **2023**, *17*, 1239069. [[CrossRef](#)]
108. Kim, J.H.; Bae, H.G.; Wu, W.C.; Nip, K.; Gould, E. SCN2A-linked myelination deficits and synaptic plasticity alterations drive auditory processing disorders in ASD. *Res. Sq.* **2024**, *10*, 3. [[CrossRef](#)]
109. Thompson, C.H.; Potet, F.; Abramova, T.V.; DeKeyser, J.M.; Ghabra, N.F.; Vanoye, C.G.; Millichap, J.J.; George, A.L. Epilepsy-associated SCN2A (NaV1.2) variants exhibit diverse and complex functional properties. *J. Gen. Physiol.* **2023**, *155*, e202313375. [[CrossRef](#)] [[PubMed](#)]
110. Fortin, J.P.; Schroeder, J.C.; Zhu, Y.; Beinborn, M.; Kopin, A.S. Pharmacological characterization of human incretin receptor missense variants. *J. Pharmacol. Exp. Ther.* **2010**, *332*, 274–280. [[CrossRef](#)] [[PubMed](#)]
111. Saxena, R.; Hivert, M.F.; Langenberg, C.; Tanaka, T.; Pankow, J.S.; Vollenweider, P.; Lyssenko, V.; Bouatia-Naji, N.; Dupuis, J.; Jackson, A.U.; et al. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. *Nat. Genet.* **2010**, *42*, 142–148. [[CrossRef](#)] [[PubMed](#)]
112. Fu, Y.; Kaneko, K.; Lin, H.Y.; Mo, Q.; Xu, Y.; Suganami, T.; Ravn, P.; Fukuda, M. Gut Hormone GIP Induces Inflammation and Insulin Resistance in the Hypothalamus. *Endocrinology* **2020**, *161*, bqaa102. [[CrossRef](#)]
113. Bekpen, C.; Kunzel, S.; Xie, C.; Eaaswarkhanth, M.; Lin, Y.L.; Gokcumen, O.; Akdis, C.A.; Tautz, D. Segmental duplications and evolutionary acquisition of UV damage response in the SPATA31 gene family of primates and humans. *BMC Genom.* **2017**, *18*, 222. [[CrossRef](#)]
114. Bekpen, C.; Xie, C.; Nebel, A.; Tautz, D. Involvement of SPATA31 copy number variable genes in human lifespan. *Aging* **2018**, *10*, 674–688. [[CrossRef](#)] [[PubMed](#)]
115. Aljabban, J.; Syed, S.; Syed, S.; Rohr, M.; Weisleder, N.; McElhanon, K.E.; Hasan, L.; Safeer, L.; Hoffman, K.; Aljabban, N.; et al. Investigating genetic drivers of dermatomyositis pathogenesis using meta-analysis. *Heliyon* **2020**, *6*, e04866. [[CrossRef](#)] [[PubMed](#)]
116. Li, S.; Wang, L.; Zhao, Q.; Wang, Z.; Lu, S.; Kang, Y.; Jin, G.; Tian, J. Genome-Wide Analysis of Cell-Free DNA Methylation Profiling for the Early Diagnosis of Pancreatic Cancer. *Front. Genet.* **2020**, *11*, 596078. [[CrossRef](#)] [[PubMed](#)]
117. Micale, L.; Fusco, C.; Augello, B.; Napolitano, L.M.; Dermitzakis, E.T.; Meroni, G.; Merla, G.; Reymond, A. Williams-Beuren syndrome TRIM50 encodes an E3 ubiquitin ligase. *Eur. J. Hum. Genet.* **2008**, *16*, 1038–1049. [[CrossRef](#)]
118. Wang, Y.; Song, W.; Zhou, S.; Chang, S.; Chang, J.; Tian, J.; Zhang, L.; Li, J.; Che, G. The genomic and transcriptome characteristics of lung adenocarcinoma patients with previous breast cancer. *BMC Cancer* **2022**, *22*, 618. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.